

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2005, 05:54:10 ; Search time 178 Seconds
(without alignments)
1527.607 Million cell updates/sec

Title: US-10-068-426-5

Perfect score: 2839

Sequence: 1 MPELLLLLLPSPHPHPIC.....MHEALHHYTKSLSPGK 531

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1615.5	56.9	626	1	GPBA_HUMAN	P07359 homo sapien
2	1222.5	43.1	481	2	O6N097	O6N097 homo sapien
3	1214.5	42.8	480	2	O6PJF1	O6PJF1 homo sapien
4	1212.5	42.7	475	2	O6GMW7	O6GMW7 homo sapien
5	1211	42.7	470	2	O6PJ44	O6PJ44 homo sapien
6	1210.5	42.6	465	2	O6GMX6	O6GMX6 homo sapien
7	1210.5	42.6	466	2	O6IN78	O6IN78 homo sapien
8	1210	42.6	476	2	O6GMX1	O6GMX1 homo sapien
9	1209.5	42.6	466	2	O6N096	O6N096 homo sapien
10	1209.5	42.6	478	2	O6PI81	O6PI81 homo sapien
11	1209	42.6	473	2	O6MZV7	O6MZV7 homo sapien
12	1208	42.6	348	2	O6PYX1	O6PYX1 homo sapien
13	1208	42.6	679	2	O96FQ8	O96FQ8 homo sapien
14	1206.5	42.5	470	2	O7Z5W1	O7Z5W1 homo sapien
15	1206	42.5	472	2	O6N089	O6N089 homo sapien
16	1205	42.4	330	1	GC1_HUMAN	P01857 homo sapien
17	1205	42.4	469	2	O7Z7P5	O7Z7P5 homo sapien
18	1204.5	42.4	475	2	O6MZQ6	O6MZQ6 homo sapien
19	1204.5	42.4	482	2	O7ZJ51	O7ZJ51 homo sapien
20	1203.5	42.4	544	2	O6PJ95	O6PJ95 homo sapien
21	1202.5	42.4	473	2	O6P055	O6P055 homo sapien
22	1201	42.3	480	2	O6N094	O6N094 homo sapien
23	1196.5	42.1	475	2	O6N095	O6N095 homo sapien
24	1181.5	41.6	487	2	O6S2L2	O6S2L2 mus sp. fv/
25	1148	40.4	326	1	GC2_HUMAN	P01859 homo sapien
26	1148	40.4	417	2	O6N093	O6N093 homo sapien
27	1145	40.3	354	2	O86TT2	O86TT2 homo sapien
28	1144	40.3	518	2	O6N030	O6N030 homo sapien
29	1143	40.3	465	2	O6P6C4	O6P6C4 homo sapien
30	1141	40.2	464	2	O6MZU6	O6MZU6 homo sapien
31	1141	40.2	521	2	O8N4Y9	O8N4Y9 homo sapien

RESULT 1

ID	GPBA_HUMAN	STANDARD;	PRT;	626 AA.
AC	P07359; Q14441; Q16469; Q8N1P3; Q8NG39; Q9HDC7; Q9UEK1; Q9UQS4;			
DT	01-APR-1988 (Rel. 07, Created)			
DT	01-APR-1988 (Rel. 07, Last sequence update)			
DT	25-OCT-2004 (Rel. 45, Last annotation update)			
DE	Platelet glycoprotein Ib alpha chain precursor (Glycoprotein Ib alpha) (GP-Ib alpha) (GP1b) (GP1b-alpha) (CD42B-alpha) (CD42B) [Contains: Glycocalicin].			
DE	Glycocalicin]			
GN	Name=GP1bA;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]_SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.			
RP	MEDLINE=87289655; PubMed=3303030;			
RX	Lopez J.A., Chung D.W., Fujikawa K., Hagen F.S., Papayannopoulou T., Roth G.J.;			
RA	"Cloning of the alpha chain of human platelet glycoprotein Ib: a transmembrane protein with homology to leucine-rich alpha 2-glycoprotein."			
RT	Proc. Natl. Acad. Sci. U.S.A. 84:5615-5619(1987).			
RL	[2]			
RN	SEQUENCE FROM N.A.			
RP	MEDLINE=89025874; PubMed=2845978;			
RX	Wenger R.H., Kieffer N., Wicki A.N., Clemetson K.J.;			
RA	"Structure of the human blood platelet membrane glycoprotein Ib alpha gene."			
RT	Biochem. Biophys. Res. Commun. 156:389-395(1988).			
RL	[3]			
RP	SEQUENCE FROM N.A., AND VARIANT PHE-86.			
RX	MEDLINE=22034223; PubMed=12038791;			
RA	Matsubara Y., Murata M., Moriki T., Yokoyama K., Watanabe N., Nakajima H., Handa M., Kawano K., Aoki K., Yoshino H., Ikeda Y.;			
RT	"A novel polymorphism, 70Leu/Phe, disrupts a consensus leu residue within the leucine-rich repeat sequence of platelet glycoprotein Ib alpha."			
RT	Thromb. Haemost. 87:867-872(2002).			
RL	[4]			
RN	SEQUENCE FROM N.A., AND VARIANT MET-161.			
RP	Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Yi Q., Nickerson D.A.;			
RA	"SeattleSNP: NHLBI HL66682 program for genomic applications, UW-FHCRC, Seattle, WA (URL: http://pga.gs.washington.edu).";			
RT	Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.			
RL	[5]			
RN	SEQUENCE FROM N.A., MUTAGENESIS OF GLY-249 AND MET-255, AND VARIANT PSEUDO-VMD SER-249.			
RP	PubMed=14521605;			
RX	Matsubara Y., Murata M., Sugita K., Ikeda Y.;			
RA	"Identification of a novel point mutation in platelet glycoprotein Ib alpha, Gly to Ser at residue 233, in a Japanese family with			

Q68CN4 homo sapien
Q8TC63 homo sapien
P01861 homo sapien
Q8NF17 homo sapien
Q6MZK7 homo sapien
P01860 homo sapien
Q35930 mus musculus
Q28256 canis famill
P01870 oryctolagus
Q95M34 equus caball
P01862 cavia porce
P22436 mus musculus
Q7Tmk1 mus musculus
Q991C4 mus musculus

ALIGNMENTS

32 1138.5 40.1 493 2 Q68CN4
33 1135 40.0 473 2 Q8TC63
34 1134.5 40.0 327 1 GC4_HUMAN
35 1130 39.8 509 2 Q8NF17
36 1127.5 39.7 476 2 Q6MZK7
37 1121.5 39.5 290 1 GC3_HUMAN
38 1032 36.4 734 2 O35930
39 927 32.7 677 2 Q28256
40 921 32.4 323 1 GC_RABIT
41 891 31.4 337 2 Q95M34
42 884.5 31.2 329 1 GC2_CAVPO
43 841 29.6 329 1 GC3_MOUSE
44 841 29.6 470 2 Q7Tmk1
45 839.5 29.6 463 2 Q991C4

RT platelet-type von Willebrand disease."; [6]
 RL J. Thromb. Haemost. 1:2198-2205 (2003).
 RN
 RP SEQUENCE FROM N.A.
 RC
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramo R.D., Mullaby S.J.,
 RA Bosak S.A., McEwen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grinwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences."; [7]
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN
 RP SEQUENCE OF 17-315.
 RC
 RX MEDLINE=87289654; PubMed=3497398;
 RA Titani K., Takio K., Handa M., Ruggeri Z.M.;
 RT "Amino acid sequence of the von Willebrand factor-binding domain of
 platelet membrane glycoprotein Ib."; [8]
 RL Proc. Natl. Acad. Sci. U.S.A. 84:5610-5614 (1987).
 RN
 RP SEQUENCE OF 305-397 FROM N.A.
 RC
 RX MEDLINE=97243178; PubMed=9088113;
 RA Suzuki K., Hayashi T., Akiba J., Yahagi A., Tajima K., Satoh S.,
 RA Sasaki H.;
 RT "StyI polymorphism at nucleotide 1610 in the human platelet
 glycoprotein Ib alpha gene."; [9]
 RL Jpn. J. Hum. Genet. 41:419-421 (1996).
 RN
 RP DISULFIDE BONDS.
 RC
 RX MEDLINE=91301149; PubMed=2070794;
 RA Hess D., Schaller J., Rickli E.E., Clemetson K.J.;
 RT "Identification of the disulphide bonds in human platelet
 glycoprotein Ib."; [10]
 RL Eur. J. Biochem. 199:389-393 (1991).
 RN
 RP INTERACTION WITH FLNB.
 RC
 RX TISSUE-Endothelial cells, and Placenta;
 RL MEDLINE=98316317; PubMed=9651345; DOI=10.1074/jbc.273.28.17531;
 RA Takafuta T., Wu G., Murphy G.F., Shapiro S.S.;
 RT "Human beta-filamin is a new protein that interacts with the
 cytoplasmic tail of glycoprotein Ibalpha."; [11]
 RL J. Biol. Chem. 273:17531-17538 (1998).
 RN
 RP X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS).
 RC
 RX MEDLINE=22171415; PubMed=12183630; DOI=10.1126/science.107355;
 RA Huizinga E.G., Teuji S., Romijn R.A., Schiphorst M.E., de Groot P.G.,
 RA Sixma J.J., Gros P.;
 RT "Structures of glycoprotein Ibalpha and its complex with von
 Willebrand factor A1 domain."; [12]
 RL Science 297:1176-1179 (2002).
 RN
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 17-304, AND SULFATION OF
 TYR-294 AND TYR-295.
 RC
 RX MEDLINE=22217968; PubMed=12087105; DOI=10.1074/jbc.M205271200;
 RA Uff S., Clemetson J.M., Harrison T., Clemetson K.J., Emsley J.;
 RT "Crystal structure of the platelet glycoprotein Ibalpha N-terminal
 domain reveals an unmasking mechanism for receptor activation."; [13]
 RL J. Biol. Chem. 277:35657-35663 (2002).
 RN
 RP VARIANT SIBA MET-161.
 RL MEDLINE=92265982; PubMed=1586750;
 RA Murata M., Furihata K., Ishida F., Russell S.R., Ware J.,
 RA Ruggeri Z.M.;
 RT "Genetic and structural characterization of an amino acid dimorphism
 in glycoprotein Ib alpha involved in platelet transfusion
 refractoriness."; [14]
 RL Blood 79:3086-3090 (1992).
 RN
 RP VARIANT BSS PHE-73.
 RL MEDLINE=92110577; PubMed=1730088;
 RA Miller J.L., Lyle V.A., Cunningham D.;
 RT "Mutation of leucine-57 to phenylalanine in a platelet glycoprotein Ib
 alpha leucine tandem repeat occurring in patients with an autosomal
 dominant variant of Bernard-Soulier disease."; [15]
 RL Blood 79:439-446 (1992).
 RN
 RP POLYMORPHISM OF PRO/THR-RICH DOMAIN.
 RL MEDLINE=92250564; PubMed=1577776;
 RA Lopez J.A., Ludwig E.H., McCarthy B.J.;
 RT "Polymorphism of human glycoprotein Ib alpha results from a variable
 number of tandem repeats of a 13-amino acid sequence in the mucin-like
 macroglycopeptide region. Structure/function implications."; [16]
 RL J. Biol. Chem. 267:10055-10061 (1992).
 RN
 RP VARIANT BSS VAL-172.
 RL MEDLINE=93388851; PubMed=7690774;
 RA Ware J., Russell S.R., Marchese P., Murata M., Mazzucato M.,
 RA de Marco L., Ruggeri Z.M.;
 RT "Point mutation in a leucine-rich repeat of platelet glycoprotein Ib
 alpha resulting in the Bernard-Soulier syndrome."; [17]
 RL J. Clin. Invest. 92:1213-1220 (1993).
 RN
 RP VARIANT BSS SER-225.
 RL MEDLINE=95118882; PubMed=7819107;
 RA Simsek S., Norris P., Lozano M., Pico M., von Dem Borne A.E.G.K.,
 RA Ribera A., Gallardo D.;
 RT "Cys209 Ser mutation in the platelet membrane glycoprotein Ib alpha
 gene is associated with Bernard-Soulier syndrome."; [18]
 RL Br. J. Haematol. 88:839-844 (1994).
 RN
 RP VARIANT PSEUDO-VMD VAL-249.
 RL MEDLINE=91271273; PubMed=2052556;
 RA Miller J.L., Cunningham D., Lyle V.A., Finch C.N.;
 RT "Mutation in the gene encoding the alpha chain of platelet
 glycoprotein Ib in platelet-type von Willebrand disease."; [19]
 RL Proc. Natl. Acad. Sci. U.S.A. 88:4761-4765 (1991).
 RN
 RP VARIANT PSEUDO-VMD VAL-249.
 RL MEDLINE=93253059; PubMed=8486780;
 RA Murata M., Russell S.R., Ruggeri Z.M., Ware J.;
 RT "Expression of the phenotypic abnormality of platelet-type von
 Willebrand disease in a recombinant glycoprotein Ib alpha fragment."; [20]
 RL J. Clin. Invest. 91:2133-2137 (1993).
 RN
 RP VARIANT PSEUDO-VMD VAL-255.
 RL MEDLINE=93214031; PubMed=8384898;
 RA Russell S.D., Roth G.J.;
 RT "Pseudo-von Willebrand disease: a mutation in the platelet
 glycoprotein Ib alpha gene associated with a hyperactive surface
 receptor."; [21]
 RL Blood 81:1787-1791 (1993).
 RN
 RP VARIANT SIBA MET-161, AND POLYMORPHISM OF PRO/THR-RICH DOMAIN.
 RL PubMed=7632942;
 RA Ishida F., Furihata K., Ishida K., Yan J., Kitano K., Kiyosawa K.,
 RA Furuta S.;
 RT "The largest variant of platelet glycoprotein Ib alpha has four tandem
 repeats of 13 amino acids in the macroglycopeptide region and a
 genetic linkage with methionine145."; [22]
 RL Blood 86:1357-1360 (1995).
 RN
 RP VARIANT BSS LEU-195 DEL.

```
RX MEDLINE=95178321; PubMed=7873390;
RA de la Salle C., Baas M.-J., Lanza F., Schwartz A., Hanau D.,
Query Match 56.9%; Score 1615.5; DB 1; Length 626;
Best Local Similarity 84.2%; Pred. No. 1.1e-102;
Matches 319; Conservative 5; Mismatches 20; Indels 35; Gaps 5;

QY 1 MPLLLLLLLLPSPHPHPICEVSKVASHLVNCDKRNLTALPPDLPKDTHLHLSNLLY 60
Db 1 MPLLLLLLLLPSPHPHPICEVSKVASHLVNCDKRNLTALPPDLPKDTHLHLSNLLY 60

QY 61 TFSLATMPYTRLTQNLDRCELTKLQVGTPLVIGTDLDSHNOLOSLPLGOTLPALT 120
Db 61 TFSLATMPYTRLTQNLDRCELTKLQVGTPLVIGTDLDSHNOLOSLPLGOTLPALT 120

QY 121 LDVSNRLTSLPLGALRGELQELYLKGNELKLTPLPGLTTPPKLEKLSANNITELP 180
Db 121 LDVSNRLTSLPLGALRGELQELYLKGNELKLTPLPGLTTPPKLEKLSANNITELP 180

QY 181 AGLLNGLENLDTLLQNSLYTTPKPGFGSHLLPPFALHGNPWLNCNCEILYFRWLQDNA 240
Db 181 AGLLNGLENLDTLLQNSLYTTPKPGFGSHLLPPFALHGNPWLNCNCEILYFRWLQDNA 240

QY 241 ENVYVWKQVVDKAVTSNVASVOCNDSKDPVVKYKPGKGCPTLGDGDTLDYDYPEEDT 300
Db 241 ENVYVWKQVVDKAVTSNVASVOCNDSKDPVVKYKPGKGCPTLGDGDTLDYDYPEEDT 300

QY 301 EGDQVR-----PHTCP-----PCPA-----PEALGAPSVFLFPK-- 330
Db 301 EGDQVR-----PHTCP-----PCPA-----PEALGAPSVFLFPK-- 330

QY 331 PKDTL-----MISRTPEVT 344
Db 361 PNFTLHMSITFSKTPKST 379

RESULT 2
Q6N097 PRELIMINARY; PRT; 481 AA.
AC Q6N097;
DT 05-JUL-2004 (T-EMBLrel. 27, Created)
DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
DE Hypothetical protein.
GN Name=DKFZp686H20196;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human esophagus tumor;
RG The German Human cDNA Consortium;
RA Wambutt R., Heubner D., Mewes H.W., Weil B., Amid C., Oeanger A.,
RA Fobo G., Han M., Wilemann S.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: BX640619; CAB45773.1;
DR HSSP: P01861; IADQ
DR InterPro: IPR003599; Ig.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003597; Ig cl.
DR InterPro: IPR003006; Ig MHC.
DR InterPro: IPR003596; Ig v.
DR Pfam: PF07654; Cl-set; 3.
DR SMART: SM00409; IG; 2.
DR SMART: SM00407; IGcl; 3.
DR SMART: SM00406; IGv; 1.
DR PROSITE: PS00835; IG LIKE; 4.
DR PROSITE: PS00290; IG_MHC;
KW Hypothetical protein.
SQ SEQUENCE 481 AA; 52759 MW; 47220D9B64BDF98B CRC64;

Query Match 43.1%; Score 1222.5; DB 2; Length 481;

Best Local Similarity 62.3%; Pred. No. 8.5e-76;
Matches 253; Conservative 13; Mismatches 53; Indels 87; Gaps 8;

QY 192 TLLLOENSLYT-----IPKFGFGSHLLPPFALHGNPWLNCNCEILYFR 233
Db 97 TLFLOWNSLRTNDTATYYCAKENSAGLLDLSYFGWKTVFAP----- 140

QY 234 RWLQDNAENY-----VWKQVVDKAVTSNVASVOCNDSKDP----- 271
Db 141 -WQGGTDVSVYASTKPGSVFPLAPSSKTSKSGTAAALGCLVKDYFPEPVTVSNWNGALTS 199

QY 272 -VYKYKGCPTLGDG-----DLDLY-----DYPEDTGDQVRP----- 307
Db 200 GVHTFPA-----VLQSSGLYSLSVVTPSSSLGTQTYICNVNHPKNTKVDKVKPEKSCD 255

QY 308 --HTCCPCPAPEALGAPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVFKNVYVDG 365
Db 256 KHTCCPCPAPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVFKNVYVDG 315

QY 366 VEVHNAKTPREBQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTSKAG 425
Db 316 VEVHNAKTPREBQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTSKAG 375

QY 426 QPREPOVYITLPSRREMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTTPVLDSD 485
Db 376 QPREPOVYITLPSRREMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTTPVLDSD 435

QY 486 GSFFLYSKLTVDKSRWQOGNVFSCSVHHEALHNNHYTKSLSPGK 531
Db 436 GSFFLYSKLTVDKSRWQOGNVFSCSVHHEALHNNHYTKSLSPGK 481

RESULT 3
Q6PJF1 PRELIMINARY; PRT; 480 AA.
AC Q6PJF1;
DT 05-JUL-2004 (T-EMBLrel. 27, Created)
DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
DE Hypothetical protein.
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.J., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
```

DR EMBL; BC016381; AAH16381.1; -
 DR HSP; P01861; IADO.
 DR InterPro; IPR003599; IG-like.
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR003597; IG CI.
 DR InterPro; IPR003006; IG MHC.
 DR InterPro; IPR003596; IG v.
 DR Pfam; PF07654; CI-set; 3.
 DR SMART; SM00409; IG; 2.
 DR SMART; SM00407; IG; 1.
 DR SMART; SM00406; IG; 1.
 DR PROSITE; PS00835; IG LIKE; 4.
 DR PROSITE; PS00290; IG MHC; UNKNOWN_2.
 KW Hypothetical protein.
 SQ SEQUENCE 480 AA; 52586 MW; 64DC641AB47CD6C8 CRC64;
 Query Match 42.8%; Score 1214.5; DB 2; Length 480;
 Best Local Similarity 65.3%; Pred. No. 3e-75;
 Matches 243; Conservative 15; Mismatches 47; Indels 67; Gaps 7;
 QY 223 WLCNCEILYFRWLQDNAENVY----VWKQVVDV-----KAVTSNV 259
 Db 113 YYCARDALYELWGSFHTDEKYGLDVGQGTPTVVSASTKGPSVFPPLAPSSKSTSGGT 172
 QY 260 ASVQCDNSDKFP-----VYKPGKGCPTLGDG-----DT 289
 Db 173 AALGCLVKDYPPEPTVVSWSNGALTSVHTFPA---VLQSSGLYSLSSVTVFPSSSLGT 228
 QY 290 DLY----DYPEEETEGDKVRP-----HTCPPCPAPEALGAPSVFLPPKPKDTLMISR 339
 Db 229 QTYICNVNHKPSNTKDKRVKFEKCDKTHTCPPCPAPELGGPSVFLPPKPKDTLMISR 288
 QY 340 TPEVTVVVDVSHEDPEVKFNWYDGVGVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNL 399
 Db 289 TPEVTVVVDVSHEDPEVKFNWYDGVGVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNL 348
 QY 400 GKEYKCKVSNKALPVPKEIKTSKAGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPS 459
 Db 349 GKEYKCKVSNKALPAPIEIKTSKAGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPS 408
 QY 460 DIAVEWESNGQPNENYKTPPVLDSGDFLYSKLTVDKSRWQGNVFPSCSWMEALHNNH 519
 Db 409 DIAVEWESNGQPNENYKTPPVLDSGDFLYSKLTVDKSRWQGNVFPSCSWMEALHNNH 468
 QY 520 YTKSLSLSLSPGK 531
 Db 469 YTKSLSLSLSPGK 480
 RESULT 4
 Q6GMW7
 ID Q6GMW7 PRELIMINARY; PRT; 475 AA.
 AC Q6GMW7;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Spleen;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Pahey J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Kzywinski M.I., Skaleka U., Smallos D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Spleen;
 RA Strausberg R.;
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC073782; AAH73782.1; -
 DR InterPro; IPR003599; IG.
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR003597; IG CI.
 DR InterPro; IPR003006; IG MHC.
 DR InterPro; IPR003596; IG v.
 DR Pfam; PF07654; CI-set; 3.
 DR SMART; SM00407; IG; 4.
 DR SMART; SM00409; IG; 2.
 DR SMART; SM00406; IG; 1.
 DR PROSITE; PS00835; IG LIKE; 4.
 DR PROSITE; PS00290; IG MHC; UNKNOWN_2.
 KW Hypothetical protein.
 SQ SEQUENCE 475 AA; 51987 MW; 2A1FE55D736860F8 CRC64;
 Query Match 42.7%; Score 1212.5; DB 2; Length 475;
 Best Local Similarity 60.6%; Pred. No. 4.1e-75;
 Matches 255; Conservative 19; Mismatches 64; Indels 83; Gaps 10;
 QY 186 GLENLDLTLQENSLY-----TIPKGFSGHLLPPAFLLHGNPWLNCILYF---R 233
 Db 63 GLEVVSIGISSENSYVADSVKGRFTISRDNKYNTL---VLQGLSLRAEDKAVYYCARA 118
 QY 234 RLQDNANVY----VWKQVVDV-----KAVTSNVASVQCDNSDKF 270
 Db 119 RCRGDTCLNFYGLDVGWQGTTVIVSSASTKGPSVFPPLAPSSKSTSGGTAALGCLVKDYF 178
 QY 271 P-----VYKPGKGCPTLGDG-----DTDIY----DYYP 296
 Db 179 PEPVTVSWSNGALTSVHTFPA---VLQSSGLYSLSSVTVFPSSSLGTQTYICNVNHK 234
 QY 297 EEDTEGDKVRP-----HTCPPCPAPEALGAPSVFLPPKPKDTLMISRPEVTCVVDV 350
 Db 235 SNTKVDKKVPEKSCDKTHTCPPCPAPELGGPSVFLPPKPKDTLMISRPEVTCVVDV 294
 QY 351 SHEDPEVKFNWYDGVGVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNK 410
 Db 295 SHEDPEVKFNWYDGVGVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNK 354
 QY 411 ALPVPKEIKTSKAGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQ 470
 Db 355 ALPAPIEIKTSKAGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQ 414
 QY 471 PENNYKTPPVLDSGDFLYSKLTVDKSRWQGNVFPSCSWMEALHNNHHTQKSLSLSPG 530
 Db 415 PENNYKTPPVLDSGDFLYSKLTVDKSRWQGNVFPSCSWMEALHNNHHTQKSLSLSPG 474
 QY 531 K 531
 Db 475 K 475
 RESULT 5
 Q6PJM4
 ID Q6PJM4 PRELIMINARY; PRT; 470 AA.

Query Match 42.6%; Score 1210.5; DB 2; Length 465;
Best Local Similarity 69.1%; Pred. No. 5.5e-75;

```
Matches 241; Conservative 11; Mismatches 45; Indels 52; Gaps 6;
QY 231 YFRWLQDQNAENV-----YVMQVVDVKAVTNSVASVQCDNSDKFP----- 271
D 121 YFDYWGQGLTVTVSSASTKGSVPFLAPSSKSTSGGTAALGCLVKDYFPEPTVSWNSGA 180
QY 272 ----VYKYPGKCGPTLGDG-----DTDLV-----DYYPEEDTEGDKVRP- 307
D 181 LTSGVHTFPA---VLQSSGLYSLSSVTVPSSSLGTYICNVNHNKPSNTKVDKKEPK 236
QY 308 -----HTCPCPAPBALGAPSVFLPPKPKDTLMISRTPEVTCVVDVSHEDDPVKKNWY 362
D 237 SCDKTHTCPCPAPBALGAPSVFLPPKPKDTLMISRTPEVTCVVDVSHEDDPVKKNWY 296
QY 363 VDGVEVHNKTKPREEQNSYRVVSVLTVLHQDLNKGKEYCKVSKNKPVPTEKTISK 422
D 297 VDGVEVHNKTKPREEQNSYRVVSVLTVLHQDLNKGKEYCKVSKNKPVPTEKTISK 356
QY 423 AKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPV 482
D 357 AKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPV 416
QY 483 DSDGSFFLYSKLTVDKSRWQOQNVFSCSVMEALHNHYTKSLSPGK 531
D 417 DSDGSFFLYSKLTVDKSRWQOQNVFSCSVMEALHNHYTKSLSPGK 465

RESULT 7
Q6IN78
ID Q6IN78 PRELIMINARY; PRT; 466 AA.
AC Q6IN78
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE IGHG1 protein.
GN Name=IGHG1;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RC TISSUE=Peripheral Nervous System;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Klausner R.D., Feingold E.A., Grouse L.H., Shermen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahay J., Helton E., Kettaman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Jones S.J., Marra M.A.;
RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RL [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Peripheral Nervous System;
RA Strausberg R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC072419; AAH72419.1;
DR HSSP; P01861; 1ADO.
DR InterPro; IPR003599; IG-
DR InterPro; IPR003597; IG-cl.
```

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalak U., Smillius D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RA Strausberg R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR ENBL; BC073773; AAH73773.1; -
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003596; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF07654; Cl-set; 3.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00409; IG; 2.
DR SMART; SM00407; IGc1; 3.
DR SMART; SM00406; IGv; 1.
DR PROSITE; PS00835; IG LIKE; 4.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_2.
KW Hypothetical protein.
SQ SEQUENCE 476 AA; 52286 MW; 622AABA5C62DDB9D CRC64;

Query Match 42.6%; Score 1210; DB 2; Length 476;
Best Local Similarity 73.7%; Pred. No. 6.1e-75;
Matches 235; Conservative 10; Mismatches 30; Indels 44; Gaps 5;

QY 253 KAVTSNVASVQCDNSDKFP-----VYKYPGKGCPTLGDG- 287
DB 162 KSTSGGTAALGCLVDFPEPTVSNVSGALTSVHTFPA----VLQSSGLSSSVTV 217

QY 288 -----DTDLY----DYPEEDTEGDKVRP-----HTCPPCPAPALGA 332
DB 218 PSSSLGTQTYICNVNHNKPSNTKVDKVEPKSCDKTHTCPCPAPELGGPSVFLPPPK 277

QY 333 DTLMIKSTPEVTCVVDVSHEDPEVKFNWYDGVVHNAKTPREQYNSTYRVSVLTV 392
DB 278 DTLMIKSTPEVTCVVDVSHEDPEVKFNWYDGVVHNAKTPREQYNSTYRVSVLTV 337

QY 393 LHQDLNGKEYKCKVSNKALPVPPIETISKAKQPREPQVYTLPPSREEMTKNOVSLTCL 452
DB 338 LHQDLNGKEYKCKVSNKALPAPIETISKAKQPREPQVYTLPPSRDELTKNOVSLTCL 397

QY 453 VKGFYPSDIAVEWESNGQPNNTKTPPVLDSGDSFFLYSKLTVDKSRWQGNVFCV 512
DB 398 VKGFYPSDIAVEWESNGQPNNTKTPPVLDSGDSFFLYSKLTVDKSRWQGNVFCV 457

QY 513 HEALHNHYTKSLSPGK 531
DB 458 HEALHNHYTKSLSPGK 476

RESULT 9
Q6N096 PRELIMINARY; PRT; 466 AA.
AC Q6N096;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein DKFZp686I15196.
GN Name=DKFZp686I15196;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human esophagus tumor;

RG The German Human cDNA Consortium;
RA Wambutt R., Heubner D., Mewes H.W., Weil B., Amid C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR ENBL; BX640620; CAE45774.1; -
DR HSSP; P01861; IADQ.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003596; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF07654; Cl-set; 3.
DR SMART; SM00409; IG; 2.
DR SMART; SM00407; IGc1; 3.
DR SMART; SM00406; IGv; 1.
DR PROSITE; PS00835; IG LIKE; 4.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_2.
KW Hypothetical protein.
SQ SEQUENCE 466 AA; 50926 MW; 01E91B748D6D57C4 CRC64;

Query Match 42.6%; Score 1209.5; DB 2; Length 466;
Best Local Similarity 62.6%; Pred. No. 6.4e-75;
Matches 244; Conservative 20; Mismatches 53; Indels 73; Gaps 7;

QY 190 LDTLLQENSLYTIKPGFPGSHLLPFAFLHGNPWLNCCEILYFRRLQDQNAENV----- 243
DB 102 LSSLRADETAVTYCARGGFGN-----FDQWGQGLTVTVSSASTK 140

QY 244 --YVMKQVVDVKAIVSNVAVOCDSNDKEP-----VYKYPGKGCPTLGDG 286
DB 141 GPSVFLAPSSKSTSGTAAALGCLVDFPEPTVSNVSGALTSVHTFPA----VLQSS 196

QY 287 G-----DTOLY----DYPEEDTEGDKVRP-----HTCPPCPAPALGA 321
DB 197 GLYSLSVVTVFSSSLGTQTYICNVNHNKPSNTKVDKVEPKSCDKTHTCPCPAPELGG 256

QY 322 PSVFLPPPKDITLMISRTPEVTCVVDVSHEDPEVKFNWYDGVVHNAKTPREQYN 381
DB 257 PSVFLPPPKDITLMISRTPEVTCVVDVSHEDPEVKFNWYDGVVHNAKTPREQYN 316

QY 382 STYRVSVLTVLHQLWNLGKEYKCKVSNKALPVPPIETISKAKQPREPQVYTLPPSRE 441
DB 317 STYRVSVLTVLHQLWNLGKEYKCKVSNKALPAPIETISKAKQPREPQVYTLPPSRDE 376

QY 442 MTKNOVSLTCLVKGFYPSDIAVEWESNGQPNNTKTPPVLDSGDSFFLYSKLTVDKSRW 501
DB 377 LTKNQVSLTCLVKGFYPSDIAVEWESNGQPNNTKTPPVLDSGDSFFLYSKLTVDKSRW 436

QY 502 QQGNVFCVSMHEALHNHYTKSLSPGK 531
DB 437 QQGNVFCVSMHEALHNHYTKSLSPGK 466

RESULT 10
Q6P181 PRELIMINARY; PRT; 478 AA.
AC Q6P181;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Primary B-Cells;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schaefer C.F., Bhat N.K.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Moore T., Wang J., Hsieh F.,
RA Hopkins R.F., Jordan H., Moore T., Wang J., Hsieh F.,

[illegible]

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
RA Diatchenko L., Maruina K.A., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan F.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grinwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzyzanski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002)."
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RA Strausberg R.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC053984; AAHS3984.1; -;
DR HSSP; P01857; 1HZH.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig cl.
DR InterPro; IPR003006; Ig MHC.
DR InterPro; IPR003596; Ig v.
DR Pfam; PF07654; C1-set; 3.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS0835; IG LIKE; 4.
DR PROSITE; PS00290; IG MHC; UNKNOWN_2.
KW Hypothetical protein.
SQ SEQUENCE 470 AA; 51204 MW; 778CF34521483E1A CRC64;

Query Match 42.5%; Score 1206.5; DB 2; Length 470;
Best Local Similarity 64.0%; Pred. No. 1e-74;
Matches 249; Conservative 15; Mismatches 64; Indels 61; Gaps 8;

Qy 191 DTLQENSLYTIKPGFGSHLLPFAFLHGNPWLNCNCEILYFRWLQDPAENV----- 243
Db DSLVLQNSLRV-----GDAVYICARGGRW---AFLGAFDINGQGTWTVSSASTKG 145

Qy 244 -YWKQVVDVKAIVTSNVASVQCDNSDKFP-----VYKPGKGCPTLGDG 287
Db PSVPLAPSSKSTSGTAAALGCLVKDYFPEPVTVSNWNSGALTSVHTFPA-----VLQSSG 201

Qy 288 -----DTDLY-----DYYPEEDTEGDKVRP-----HTCCPCPAPALGAP 322
Db LYSLSVVTVFSSSLGTQTYICNVNHPKSNKTKVDRVPEKSCDTHTCPCPAPALGAP 261

Qy 323 SVFLPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQVNS 382
Db SVFLPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQVNS 321

Qy 383 TYRVVSVLTVLHQDWLNGKEYKCKVSNKALPVPKEKTSISKAGQPREPQVYTLPPSREEM 442
Db TYRVVSVLTVLHQDWLNGKEYKCKVSNKALPVPKEKTSISKAGQPREPQVYTLPPSREEM 381

Qy 443 TKQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSGFFLYSKLTVDKSRWQ 502
Db TKQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSGFFLYSKLTVDKSRWQ 441

Qy 503 QGNVFCSCVMHEALHNHYTQKSLSLSPGK 531
Db QGNVFCSCVMHEALHNHYTQKSLSLSPGK 470

RESULT 15

Q6N089 PRELIMINARY; PRT; 472 AA.
AC Q6N089,
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein DKFZp686P15220.
GN Name=DKFZp686P15220;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human rectum tumor;
RG The German Human cDNA Consortium;
RA Wambutt R., Heubner D., Mewes H.W., Well B., Amid C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX640627; CAB45781.1; -;
DR HSSP; P01861; IADQ.
DR InterPro; IPR003599; Ig-like.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig cl.
DR InterPro; IPR003006; Ig MHC.
DR InterPro; IPR003596; Ig v.
DR Pfam; PF07654; C1-set; 3.
DR SMART; SM00409; IG; 2.
DR SMART; SM00407; IGcl; 3.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS0835; IG LIKE; 4.
DR PROSITE; PS00290; IG MHC; UNKNOWN_2.
KW Hypothetical protein.
SQ SEQUENCE 472 AA; 51724 MW; 26CB340D0046D279 CRC64;

Query Match 42.5%; Score 1206; DB 2; Length 472;
Best Local Similarity 73.4%; Pred. No. 1.1e-74;
Matches 234; Conservative 11; Mismatches 30; Indels 44; Gaps 5;

Qy 253 KAVTSNVASVQCDNSDKFP-----VYKPGKGCPTLGDG----- 287
Db 158 KSTSGGTAALGCLVKDYFPEPVTVSNWNSGALTSVHTFPA-----VLQSSGLYSLSSVTV 213

Qy 288 -----DTDLY-----DYYPEEDTEGDKVRP-----HTCCPCPAPALGAPSVFLPPKPK 332
Db PSSSLGTQTYICNVNHPKSNKTKVDRVPEKSCDTHTCPCPAPALGAPSVFLPPKPK 273

Qy 333 DTLMSRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQVNSYRVSVLTV 392
Db DTLMSRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQVNSYRVSVLTV 333

Qy 393 LHQDWLNGKEYKCKVSNKALPVPKEKTSISKAGQPREPQVYTLPPSREEMTKNOVSLTCL 452
Db LHQDWLNGKEYKCKVSNKALPVPKEKTSISKAGQPREPQVYTLPPSREEMTKNOVSLTCL 393

Qy 453 VKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSGFFLYSKLTVDKSRWQGNVFCSCVM 512
Db VKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSGFFLYSKLTVDKSRWQGNVFCSCVM 453

Qy 513 HEALHNHYTQKSLSLSPGK 531
Db HEALHNHYTQKSLSLSPGK 472

Search completed: May 24, 2005, 05:57:19
Job time : 181 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2005, 05:54:10 ; Search time 162 Seconds
(without alignments)
1267.715 Million cell updates/sec

Title: US-10-068-426-5
Perfect score: 2839
Sequence: 1 MPELLLLLLLPLPLPHPIC.....MHEALHNHYTKSLSPGK 531

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_16Dec04:.*
1: Geneseqp1980s:.*
2: Geneseqp1990s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003as:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2839	100.0	531	5	Abb78238 Amino aci
2	2829	99.6	531	5	Abb78237 Amino aci
3	2822.5	99.4	544	5	Abb78236 Amino aci
4	2822	99.4	531	5	Abb78239 Amino aci
5	2812.5	99.1	544	5	Abb78234 Amino aci
6	2812.5	99.1	544	5	Abb78235 Amino aci
7	2411.5	84.9	562	3	Aay49933 Human gly
8	2381.5	83.9	568	3	Aay49935 Human gly
9	1615.5	56.9	626	4	Aae12135 Human gly
10	1615.5	56.9	626	7	Adf69094 Human Mp5
11	1615.5	56.9	626	7	Adj68624 Human hea
12	1615.5	56.9	626	8	Adh40324 Human pla
13	1615.5	56.9	626	8	Adq39833 Human myo
14	1536	54.1	290	5	Abb78244 Amino aci
15	1527	53.8	301	5	Abb78240 Amino aci
16	1526	53.8	290	5	Abb78243 Amino aci
17	1526	53.8	233	1	Aap91368 45 kDa am
18	1521.5	53.6	302	5	Abb78241 Amino aci
19	1521.5	53.6	610	2	Aar89436 Mutated p
20	1521.5	53.6	610	2	Aar18201 Platelet
21	1520	53.5	290	5	Abb78245 Amino aci
22	1518.5	53.5	610	2	Aar51116 Platelet
23	1514.5	53.3	610	2	Aar56664 Mutated pl
24	1513.5	53.3	300	5	Abb78242 Amino aci
25	1372	48.3	609	8	Adp76033 AMIGO pol

26	1357.5	47.8	640	8	ADP76035	Adp76035 AMIGO pol
27	1324	46.6	629	8	ADP76037	Adp76037 AMIGO pol
28	1279	45.1	800	7	ADP03635	Ade03635 BGS assoc
29	1237	43.6	313	2	AAW53321	Aaw53321 P-selecti
30	1237	43.6	313	2	AAV29766	Aay29766 P-selecti
31	1237	43.6	313	5	AAg77946	Aag77946 Human dim
32	1237	43.6	313	6	ABR39961	Abr39961 rFSGI-Ig
33	1235.5	43.5	420	5	AAE15348	Aae15348 Human ery
34	1234	43.5	778	2	AAW15489	Aaw15489 Pemphigus
35	1232.5	43.4	470	6	ABR61415	Abr61415 Human IL-
36	1231	43.4	581	8	ADP03590	Adp03590 Infection
37	1230.5	43.3	401	2	AAW10537	Aaw10537 Leptin 1-
38	1230	43.3	388	4	AAW50248	Aam50248 Human int
39	1230	43.3	388	8	ADQ76791	Adq76791 KIN-2, an
40	1230	43.3	449	2	AAW96278	Aaw96278 Human nog
41	1230	43.3	502	6	ABJ37109	Abj37109 Concatame
42	1230	43.3	502	8	ADQ79922	Adq79922 Human mgC
43	1229.5	43.3	542	5	AAU76357	Aau76357 Pc disint
44	1228.5	43.3	528	5	AAU79654	Aau79654 Human sol
45	1228.5	43.3	528	8	ADN59105	Adn59105 Human sol

ALIGNMENTS

RESULT 1					
ABE78238					
ID	ABB78238	standard; protein; 531 AA.			
XX	AC	ABB78238;			
XX	DT	25-NOV-2002 (first entry)			
XX	DE	Amino acid sequence of GPLb290/2V-Ig fusion protein.			
XX	KW	Glycoprotein 1B-alpha; GPLb; immunoglobulin; Ig; platelet adherence;			
XX	KW	leukocyte; platelet activation; ischaemic heart disease;			
XX	KW	acute myocardial infarction; stroke; venous thrombosis; atherosclerosis;			
XX	KW	arterial thrombosis; angina; vascular condition; vascular inflammation;			
XX	OS	thrombosis; angioplasty; restenosis.			
XX	PH	Synthetic.			
FT	Key	Location/Qualifiers			
FT	Misc-difference	487			
FT	/note= "Ser encoded by CCC"				
XX	PN	WO200263003-A2.			
XX	PD	15-AUG-2002.			
XX	PF	06-FEB-2002; 2002WO-US003549.			
XX	PR	06-FEB-2001; 2001US-0266838P.			
XX	PA	(GEMY) GENETICS INST LLC.			
XX	PI	Shaw GD, Kumar R, Sako DS, Mcdonagh T, Sullivan FX;			
XX	DR	WPI; 2002-657537/70.			
XX	DR	N-PSDB; ABQ78665.			
XX	PT	New glycoprotein 1b alpha fusion polypeptides, useful for treating a			
XX	PT	disorder associated with platelet activation e.g. ischemic heart disease,			
XX	PT	stroke, venous or arterial thrombosis or atherosclerosis.			
XX	PS	Claim 20; Page 3-4; 45pp; English.			
XX	CC	The present sequence represents a fusion protein of glycoprotein 1B-alpha			
XX	CC	(GPLb) and an immunoglobulin (Ig) polypeptide. The fusion protein			
XX	CC	inhibits the adherence of platelets to leukocytes. The fusion polypeptide			
XX	CC	is useful for treating a disorder associated with platelet activation			
XX	CC	e.g. ischaemic heart disease, acute myocardial infarction, stroke, venous			

CC thrombosis, atherosclerosis, arterial thrombosis or unstable angina. It
CC can also be used to treat vascular conditions associated with vascular
CC inflammation, thrombosis, and angioplasty-related restenosis
XX
SQ Sequence 531 AA;

Query Match 100.0%; Score 2839; DB 5; Length 531;
Best Local Similarity 100.0%; Pred. No. 6.7e-205;
Matches 531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MPLLILLLLLPSPLPHPHICVSKVASHLEVNCNRLTALPPDPKDTTILHLENLLY 60
Db 1 MPLLILLLLLPSPLPHPHICVSKVASHLEVNCNRLTALPPDPKDTTILHLENLLY 60
Qy 61 TFSLATLMPYTRLTQNLDRCELTKLQVDGTLPLVLTDLSHNQLQSLPLAGTLPALT 120
Db 61 TFSLATLMPYTRLTQNLDRCELTKLQVDGTLPLVLTDLSHNQLQSLPLAGTLPALT 120
Qy 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLP 180
Db 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLP 180
Qy 181 AGLNGLENLDTLLQENSLYTIIPKFFGSHLLPFAFLHGNPWLNCNCEILYFRWLQNA 240
Db 181 AGLNGLENLDTLLQENSLYTIIPKFFGSHLLPFAFLHGNPWLNCNCEILYFRWLQNA 240
Qy 241 ENVYVWKQVDVKAVTSNVASVQCDNSDKFPVYKPGKCPGLDGDGTDLYDYPEDT 300
Db 241 ENVYVWKQVDVKAVTSNVASVQCDNSDKFPVYKPGKCPGLDGDGTDLYDYPEDT 300
Qy 301 EGDVVRHTCCPCPAPEALGAPSVFLPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFN 360
Db 301 EGDVVRHTCCPCPAPEALGAPSVFLPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFN 360
Qy 361 WYVDGVEVHNAKTPREEQYNSTYRVSVLTVLHQDWLNGKEYCKVSNKALPVPKEITI 420
Db 361 WYVDGVEVHNAKTPREEQYNSTYRVSVLTVLHQDWLNGKEYCKVSNKALPVPKEITI 420
Qy 421 SKAQGPPEQVYTLPPSREEMTKNQVSLTCLVKGFPSPDIADVWESNGQPNYKTTTP 480
Db 421 SKAQGPPEQVYTLPPSREEMTKNQVSLTCLVKGFPSPDIADVWESNGQPNYKTTTP 480
Qy 481 VLDSGSPFLYSKLTVDKSRWQGNVFCSCVMHEALHNNHYTKQSLSPGK 531
Db 481 VLDSGSPFLYSKLTVDKSRWQGNVFCSCVMHEALHNNHYTKQSLSPGK 531

RESULT 2
ID ABB78237
XX ABB78237 standard; protein; 531 AA.

AC ABB78237;

XX 25-NOV-2002 (first entry)

DE Amino acid sequence of GPIIb290-Ig fusion protein.

XX Glycoprotein IIB-alpha; GPIb; immunoglobulin; Ig; platelet adherence;
KW leukocyte; platelet activation; ischaemic heart disease;
KW acute myocardial infarction; stroke; venous thrombosis; atherosclerosis;
KW arterial thrombosis; angina; vascular condition; vascular inflammation;
KW thrombosis; angioplasty; restenosis.

OS Synthetic.

XX Key Location/Qualifiers
FH Misc-difference 487
FT /note= "Ser encoded by CCC"

XX WO200263003-A2.

PD 15-AUG-2002.

XX

PF 06-FEB-2002; 2002WO-US003549.
XX
PR 06-FEB-2001; 2001US-0266838P.
XX
XX (GEMY) GENETICS INST LLC.
XX
XX Shaw GD, Kumar R, Sako DS, Mcdonagh T, Sullivan FX;
XX WPI: 2002-657537/70.
XX N-PSDB; ABQ78664.
XX
XX New glycoprotein Iib alpha fusion polypeptides, useful for treating a
XX disorder associated with platelet activation e.g. ischemic heart disease,
XX stroke, venous or arterial thrombosis or atherosclerosis.
XX
XX Claim 20; Page 3; 45pp; English.
XX
XX The present sequence represents a fusion protein of glycoprotein IIB-alpha
XX (GPIb) and an immunoglobulin (Ig) polypeptide. The fusion protein
XX inhibits the adherence of platelets to leukocytes. The fusion polypeptide
XX is useful for treating a disorder associated with platelet activation
XX e.g. ischaemic heart disease, acute myocardial infarction, stroke, venous
XX thrombosis, atherosclerosis, arterial thrombosis or unstable angina. It
XX can also be used to treat vascular conditions associated with vascular
XX inflammation, thrombosis, and angioplasty-related restenosis
XX
SQ Sequence 531 AA;

Query Match 99.6%; Score 2829; DB 5; Length 531;
Best Local Similarity 99.6%; Pred. No. 3.8e-204;
Matches 529; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MPLLILLLLLPSPLPHPHICVSKVASHLEVNCNRLTALPPDPKDTTILHLENLLY 60
Db 1 MPLLILLLLLPSPLPHPHICVSKVASHLEVNCNRLTALPPDPKDTTILHLENLLY 60
Qy 61 TFSLATLMPYTRLTQNLDRCELTKLQVDGTLPLVLTDLSHNQLQSLPLAGTLPALT 120
Db 61 TFSLATLMPYTRLTQNLDRCELTKLQVDGTLPLVLTDLSHNQLQSLPLAGTLPALT 120
Qy 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLP 180
Db 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLP 180
Qy 181 AGLNGLENLDTLLQENSLYTIIPKFFGSHLLPFAFLHGNPWLNCNCEILYFRWLQNA 240
Db 181 AGLNGLENLDTLLQENSLYTIIPKFFGSHLLPFAFLHGNPWLNCNCEILYFRWLQNA 240
Qy 241 ENVYVWKQVDVKAVTSNVASVQCDNSDKFPVYKPGKCPGLDGDGTDLYDYPEDT 300
Db 241 ENVYVWKQVDVKAVTSNVASVQCDNSDKFPVYKPGKCPGLDGDGTDLYDYPEDT 300
Qy 301 EGDVVRHTCCPCPAPEALGAPSVFLPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFN 360
Db 301 EGDVVRHTCCPCPAPEALGAPSVFLPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFN 360
Qy 361 WYVDGVEVHNAKTPREEQYNSTYRVSVLTVLHQDWLNGKEYCKVSNKALPVPKEITI 420
Db 361 WYVDGVEVHNAKTPREEQYNSTYRVSVLTVLHQDWLNGKEYCKVSNKALPVPKEITI 420
Qy 421 SKAQGPPEQVYTLPPSREEMTKNQVSLTCLVKGFPSPDIADVWESNGQPNYKTTTP 480
Db 421 SKAQGPPEQVYTLPPSREEMTKNQVSLTCLVKGFPSPDIADVWESNGQPNYKTTTP 480
Qy 481 VLDSGSPFLYSKLTVDKSRWQGNVFCSCVMHEALHNNHYTKQSLSPGK 531
Db 481 VLDSGSPFLYSKLTVDKSRWQGNVFCSCVMHEALHNNHYTKQSLSPGK 531

RESULT 3
ABB78236
ID ABB78236 standard; protein; 544 AA.
XX

AC ABB78236;
 XX 25-NOV-2002 (first entry)
 DT Amino acid sequence of GPlb302/4X-Ig fusion protein.
 DE
 XX Glycoprotein 1B-alpha; GPlb; immunoglobulin; Ig; platelet adherence;
 KW leukocyte; platelet activation; ischaemic heart disease;
 KW acute myocardial infarction; stroke; venous thrombosis; atherosclerosis;
 KW arterial thrombosis; angina; vascular condition; vascular inflammation;
 KW thrombosis; angioplasty; restenosis.
 XX
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FH Misc-difference 500 /note= "Ser encoded by CCC"
 FT
 PN WO200263003-A2.
 XX
 XX 15-AUG-2002.
 PD
 XX 06-FEB-2002; 2002WO-US003549.
 XX
 XX 06-FEB-2001; 2001US-0266838P.
 PR
 XX (GEMY) GENETICS INST LLC.
 PA
 XX Shaw GD, Kumar R, Sako DS, Mcdonagh T, Sullivan FX;
 PI
 XX WPI; 2002-657537/70.
 DR N-PSDB; ABQ78663.
 XX
 XX New glycoprotein 1b alpha fusion polypeptides, useful for treating a
 PT disorder associated with platelet activation e.g. ischemic heart disease,
 PT stroke, venous or arterial thrombosis or atherosclerosis.
 XX
 PS Claim 20; Page 3; 45pp; English.
 XX
 XX The present sequence represents a fusion protein of glycoprotein 1B-alpha
 CC (GPlb) and an immunoglobulin (Ig) polypeptide. The fusion protein
 CC inhibits the adherence of platelets to leukocytes. The fusion polypeptide
 CC is useful for treating a disorder associated with platelet activation
 CC e.g. ischaemic heart disease, acute myocardial infarction, stroke, venous
 CC thrombosis, atherosclerosis, arterial thrombosis or unstable angina. It
 CC can also be used to treat vascular conditions associated with vascular
 CC inflammation, thrombosis, and angioplasty-related restenosis
 XX
 SQ Sequence 544 AA;
 Query Match 99.4%; Score 2822.5; DB 5; Length 544;
 Best Local Similarity 97.6%; Pred. No. 1.2e-203;
 Matches 531; Conservative 0; Mismatches 0; Indels 13; Gaps 1;
 QY 1 MPLLLLLLLLPSPHPHPICEVSKVASHLEVNCCKRNLTALPPDLPKOTTLHLSENLLY 60
 DB 1 MPLLLLLLLLPSPHPHPICEVSKVASHLEVNCCKRNLTALPPDLPKOTTLHLSENLLY 60
 QY 61 TFSLATMPYTRLTQNLNRCCLTKLOVDGTLPLVGLTDLSHNQLSLPLGOTLPALTV 120
 DB 61 TFSLATMPYTRLTQNLNRCCLTKLOVDGTLPLVGLTDLSHNQLSLPLGOTLPALTV 120
 QY 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLP 180
 DB 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLP 180
 QY 181 AGLLNGLENLDTLLQENSLYTIPIKGFSGSHLLPFAFLHGNPWLNCCEILYFRRWLQDNA 240
 DB 181 AGLLNGLENLDTLLQENSLYTIPIKGFSGSHLLPFAFLHGNPWLNCCEILYFRRWLQDNA 240
 QY 241 ENYVWKQVVDKAVTNSVAVQCDNSDKFVYKYGKCPILGDEGDTLDYDYPEEDT 300
 DB 241 ENYVWKQVVDKAVTNSVAVQCDNSDKFVYKYGKCPILGDEGDTLDYDYPEEDT 300

QY 301 EGDKV-----RPHTCPPCPAPEALGASVFLFPKPKDITLMISRTPEVTCVV 347
 DB 301 EGDKVAATATVVKFTKARPHTCPPCPAPEALGASVFLFPKPKDITLMISRTPEVTCVV 360
 QY 348 VDVSHEDPEVKENWYVDGVEVHNATKPREOYNSTYRVSVLTVLHODMLNGKEYCKV 407
 DB 361 VDVSHEDPEVKENWYVDGVEVHNATKPREOYNSTYRVSVLTVLHODMLNGKEYCKV 420
 QY 408 SNKALPVPTEKTSKAKGQPREPQVYVTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEMES 467
 DB 421 SNKALPVPTEKTSKAKGQPREPQVYVTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEMES 480
 QY 468 NQOPENNYKTTTPVLDSDSGSFPLYSKLTVDKSRWQGNVFCVSMHEALHNYTQKSLSL 527
 DB 481 NQOPENNYKTTTPVLDSDSGSFPLYSKLTVDKSRWQGNVFCVSMHEALHNYTQKSLSL 540
 QY 528 SPGK 531
 DB 541 SPGK 544
 RESULT 4
 ABB78239
 ID ABB78239 standard; protein; 531 AA.
 XX
 AC ABB78239;
 DT 25-NOV-2002 (first entry)
 XX
 DE Amino acid sequence of GPlb290/1A-Ig fusion protein.
 XX
 KW Glycoprotein 1B-alpha; GPlb; immunoglobulin; Ig; platelet adherence;
 KW leukocyte; platelet activation; ischaemic heart disease;
 KW acute myocardial infarction; stroke; venous thrombosis; atherosclerosis;
 KW arterial thrombosis; angina; vascular condition; vascular inflammation;
 KW thrombosis; angioplasty; restenosis.
 XX
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FT Misc-difference 253 /note= "Val encoded by GCG"
 FT Misc-difference 487 /note= "Ser encoded by CCC"
 XX
 PN WO200263003-A2.
 XX
 PD 15-AUG-2002.
 XX
 XX 06-FEB-2002; 2002WO-US003549.
 XX
 XX 06-FEB-2001; 2001US-0266838P.
 PR
 XX (GEMY) GENETICS INST LLC.
 PA
 XX Shaw GD, Kumar R, Sako DS, Mcdonagh T, Sullivan FX;
 PI
 XX WPI; 2002-657537/70.
 DR N-PSDB; ABQ78666.
 XX
 XX New glycoprotein 1b alpha fusion polypeptides, useful for treating a
 PT disorder associated with platelet activation e.g. ischemic heart disease,
 PT stroke, venous or arterial thrombosis or atherosclerosis.
 XX
 PS Claim 20; Page 4; 45pp; English.
 XX
 XX The present sequence represents a fusion protein of glycoprotein 1B-alpha
 CC (GPlb) and an immunoglobulin (Ig) polypeptide. The fusion protein
 CC inhibits the adherence of platelets to leukocytes. The fusion polypeptide
 CC is useful for treating a disorder associated with platelet activation
 CC e.g. ischaemic heart disease, acute myocardial infarction, stroke, venous
 CC thrombosis, atherosclerosis, arterial thrombosis or unstable angina. It

CC can also be used to treat vascular conditions associated with vascular
CC inflammation, thrombosis, and angioplasty-related restenosis
XX
XX
SQ Sequence 531 AA;

Query Match 99.4%; Score 2822; DB 5; Length 531;
Best Local Similarity 99.4%; Pred. No. 1.3e-203;
Matches 528; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MPLLALLLPSPLPHPHICVSKVASHLEVNCNRLTALPPDLPKDTTILHSENLLY 60
Dy 1 MPLLALLLPSPLPHPHICVSKVASHLEVNCNRLTALPPDLPKDTTILHSENLLY 60
Qy 61 TFSLATLMPYTRLTQNLDRCELTKLQVDGTLPLVLTDLDSHNOLQSLPLGOTLPALT 120
Dy 61 TFSLATLMPYTRLTQNLDRCELTKLQVDGTLPLVLTDLDSHNOLQSLPLGOTLPALT 120
Qy 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLP 180
Dy 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLP 180
Qy 181 AGLNGLENLDTLLQENSLYTI PKGFFGSHLLPFAPLHGNPWLNCCEILYFRRWLQDNA 240
Dy 181 AGLNGLENLDTLLQENSLYTI PKGFFGSHLLPFAPLHGNPWLNCCEILYFRRWLQDNA 240
Qy 241 ENVYVWKQVDVKAITSNVASVQCDNSDKFPVYKPGKGCPTLGDGDTLDLYYPEEDT 300
Dy 241 ENVYVWKQVDVKAITSNVASVQCDNSDKFPVYKPGKGCPTLGDGDTLDLYYPEEDT 300
Qy 301 EGDVVRHTCPCPAPEALGAPSVFLFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFN 360
Dy 301 EGDVVRHTCPCPAPEALGAPSVFLFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFN 360
Qy 361 WYVDGVEVHNAKTPREEQYNSTRVSVLTVLHODWLNGKEYCKVSNKALPVPIEKT 420
Dy 361 WYVDGVEVHNAKTPREEQYNSTRVSVLTVLHODWLNGKEYCKVSNKALPVPIEKT 420
Qy 421 SKAKGQPREPOVYTLPPSREEMTKNQVSLTCLVKGFPYPSDIAVWESNGQPENNYKTTP 480
Dy 421 SKAKGQPREPOVYTLPPSREEMTKNQVSLTCLVKGFPYPSDIAVWESNGQPENNYKTTP 480
Qy 481 VLDSGSPFLYSKLTVDKSRWQGNVFCVSWHEALHNNHYTKQSLSPGK 531
Dy 481 VLDSGSPFLYSKLTVDKSRWQGNVFCVSWHEALHNNHYTKQSLSPGK 531

RESULT 5
ABB78234 standard; protein; 544 AA.
XX
XX ABB78234;
XX
XX 25-NOV-2002 (first entry)
XX
XX Amino acid sequence of GPIb302-Ig fusion protein.
XX
XX Glycoprotein 1B-alpha; GPIb; immunoglobulin; Ig; platelet adherence;
KW leukocyte; platelet activation; ischaemic heart disease;
KW acute myocardial infarction; stroke; venous thrombosis; atherosclerosis;
KW arterial thrombosis; angina; vascular condition; vascular inflammation;
KW thrombosis; angioplasty; restenosis.
XX
XX Synthetic.
XX
XX
XX Key Location/Qualifiers
XX Misc-difference 500 /note= "Ser encoded by CCC"
XX
XX WO200263003-A2.
XX
XX 15-AUG-2002.
XX
XX 06-FEB-2002; 2002WO-US003549.

XX 06-FEB-2001; 2001US-0266838P.
XX (GEMY) GENETICS INST LLC.
XX
XX Shaw GD, Kumar R, Sako DS, Mcdonagh T, Sullivan FX;
XX WPI; 2002-657537/70.
XX N-PSDB; ABQ78661.
XX
XX New glycoprotein 1b alpha fusion polypeptides, useful for treating a
XX disorder associated with platelet activation e.g. ischemic heart disease,
XX stroke, venous or arterial thrombosis or atherosclerosis.
XX
XX Claim 5; Page 2; 45pp; English.
XX
XX The present sequence represents a fusion protein of glycoprotein 1B-alpha
XX (GPIb) and an immunoglobulin (Ig) polypeptide. The fusion protein
XX inhibits the adherence of platelets to leukocytes. The fusion polypeptide
XX is useful for treating a disorder associated with platelet activation
XX e.g. ischaemic heart disease, acute myocardial infarction, stroke, venous
XX thrombosis, atherosclerosis, arterial thrombosis or unstable angina. It
XX can also be used to treat vascular conditions associated with vascular
XX inflammation, thrombosis, and angioplasty-related restenosis
XX
SQ Sequence 544 AA;

Query Match 99.1%; Score 2812.5; DB 5; Length 544;
Best Local Similarity 97.2%; Pred. No. 6.8e-203;
Matches 529; Conservative 1; Mismatches 1; Indels 13; Gaps 1;

Qy 1 MPLLALLLPSPLPHPHICVSKVASHLEVNCNRLTALPPDLPKDTTILHSENLLY 60
Dy 1 MPLLALLLPSPLPHPHICVSKVASHLEVNCNRLTALPPDLPKDTTILHSENLLY 60
Qy 61 TFSLATLMPYTRLTQNLDRCELTKLQVDGTLPLVLTDLDSHNOLQSLPLGOTLPALT 120
Dy 61 TFSLATLMPYTRLTQNLDRCELTKLQVDGTLPLVLTDLDSHNOLQSLPLGOTLPALT 120
Qy 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLP 180
Dy 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLP 180
Qy 181 AGLNGLENLDTLLQENSLYTI PKGFFGSHLLPFAPLHGNPWLNCCEILYFRRWLQDNA 240
Dy 181 AGLNGLENLDTLLQENSLYTI PKGFFGSHLLPFAPLHGNPWLNCCEILYFRRWLQDNA 240
Qy 241 ENVYVWKQVDVKAITSNVASVQCDNSDKFPVYKPGKGCPTLGDGDTLDLYYPEEDT 300
Dy 241 ENVYVWKQVDVKAITSNVASVQCDNSDKFPVYKPGKGCPTLGDGDTLDLYYPEEDT 300
Qy 301 EGDVVRHTCPCPAPEALGAPSVFLFPKPKDTLMISRTPEVTCV 347
Dy 301 EGDVVRHTCPCPAPEALGAPSVFLFPKPKDTLMISRTPEVTCV 360
Qy 348 VDVSHEDPEVKFNWYVDGVEVHNAKTPREEQYNSTRVSVLTVLHODWLNGKEYCKV 407
Dy 361 VDVSHEDPEVKFNWYVDGVEVHNAKTPREEQYNSTRVSVLTVLHODWLNGKEYCKV 420
Qy 408 SNKALPVPIEKTISKAKGQPREPOVYTLPPSREEMTKNQVSLTCLVKGFPYPSDIAVWES 467
Dy 421 SNKALPVPIEKTISKAKGQPREPOVYTLPPSREEMTKNQVSLTCLVKGFPYPSDIAVWES 480
Qy 468 NGQPENNYKTTPPVLDSDGSPFLYSKLTVDKSRWQGNVFCVSWHEALHNNHYTKQSL 527
Dy 481 NGQPENNYKTTPPVLDSDGSPFLYSKLTVDKSRWQGNVFCVSWHEALHNNHYTKQSL 540
Qy 528 SPGK 531
Dy 541 SPGK 544

RESULT 6

```

ABB78235
ID ABB78235 standard; protein; 544 AA.
XX
AC ABB78235;
XX
XX 25-NOV-2002 (first entry)
XX
DE Amino acid sequence of GPIb302/2A-Ig fusion protein.
XX
XX Glycoprotein Ib-alpha; GPIb; immunoglobulin; Ig; platelet adherence;
XX leukocyte; platelet activation; ischaemic heart disease;
XX acute myocardial infarction; stroke; venous thrombosis; atherosclerosis;
XX arterial thrombosis; angina; vascular condition; vascular inflammation;
XX thrombosis; angioplasty; restenosis.
XX
OS Synthetic.
XX
XX Key Location/Qualifiers
XX Key-difference 500
XX Misc-difference 500 /note= "Ser encoded by CCC"
XX
XX WO200263003-A2.
XX
XX 15-AUG-2002.
XX
XX 06-FEB-2002; 2002WO-US0003549.
XX
XX 06-FEB-2001; 2001US-0266838P.
XX
XX (GEMY ) GENETICS INST LLC.
XX
XX Shaw GD, Kumar R, Sako DS, McDonagh T, Sullivan FX;
XX
XX WPI; 2002-657537/70.
XX
XX N-PSDB; ABQ78662.
XX
XX New glycoprotein Ib alpha fusion polypeptides, useful for treating a
XX disorder associated with platelet activation e.g. ischemic heart disease,
XX stroke, venous or arterial thrombosis or atherosclerosis.
XX
XX Claim 20; Page 3; 45pp; English.
XX
XX The present sequence represents a fusion protein of glycoprotein Ib-alpha
XX (GPIb) and an immunoglobulin (Ig) polypeptide. The fusion protein
XX inhibits the adherence of platelets to leukocytes. The fusion polypeptide
XX is useful for treating a disorder associated with platelet activation
XX e.g. ischaemic heart disease, acute myocardial infarction, stroke, venous
XX thrombosis, atherosclerosis, arterial thrombosis or unstable angina. It
XX can also be used to treat vascular conditions associated with vascular
XX inflammation, thrombosis, and angioplasty-related restenosis
XX
XX Sequence 544 AA;
XX
Query Match 99.1%; Score 2812.5; DB 5; Length 544;
Best Local Similarity 97.2%; Pred. No. 6.8e-203;
Matches 529; Conservative 1; Mismatches 1; Indels 13; Gaps 1;
XX
XX 1 MPLLLLLLLLPSLPHPHPICEVSKVASHLEVNCDEKNTALPDLPKDTTLHLSENILY 60
XX
XX 1 MPLLLLLLLLPSLPHPHPICEVSKVASHLEVNCDEKNTALPDLPKDTTLHLSENILY 60
XX
XX 61 TFSLATLMPYTRLTQNLDRCELTKLQVDTGLPVLTGLDLSHNLQSLPLLGOTLPALT 120
XX
XX 61 TFSLATLMPYTRLTQNLDRCELTKLQVDTGLPVLTGLDLSHNLQSLPLLGOTLPALT 120
XX
XX 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGLTPTPKLEKLSLANNNTLELP 180
XX
XX 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGLTPTPKLEKLSLANNNTLELP 180
XX
XX 181 AGLLNGLENLDTLLQENSLYTPKGFSGSHLLPFAFLHGNPWLNCCEILYFRRWLODNA 240
XX
XX 181 AGLLNGLENLDTLLQENSLYTPKGFSGSHLLPFAFLHGNPWLNCCEILYFRRWLODNA 240
XX

```

```

QY 241 ENVYVWKQVVDVKAVTNSVASVQCDNSDKFPVYKYPGKGCPTLGDSDGDTLDLYYPEEDT 300
DB 241 ENVYVWKQVVDVKAVTNSVASVQCDNSDKFPVYKYPGKGCPTLGDSDGDTLDLYYPEEDT 300
QY 301 EGDKV-----RPHTCPCPAPENLGAHSVFLPPPKDPTLMISRTPEVTCVV 347
DB 301 EGDKVAATATVVKFPPTKARPHTCPPCPAPENLGAHSVFLPPPKDPTLMISRTPEVTCVV 360
QY 348 DVVSHEDPEVKFNWYVDGVEVHNNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKV 407
DB 361 DVVSHEDPEVKFNWYVDGVEVHNNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKV 420
QY 408 SNKALPVPTEKTSKAKGQRPSPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWES 467
DB 421 SNKALPVPTEKTSKAKGQRPSPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWES 480
QY 468 NGQPENNYKTTTPVLDSGSPFLYSLKLTVDKSRWQGNVFPSCVMHEALHNNHYTKQSLSL 527
DB 481 NGQPENNYKTTTPVLDSGSPFLYSLKLTVDKSRWQGNVFPSCVMHEALHNNHYTKQSLSL 540
QY 528 SPGK 531
DB 541 SPGK 544

RESULT 7
AA49933
ID AA49933 standard; protein; 562 AA.
XX
AC AA49933;
XX
DT 12-SEP-2003 (revised)
DT 01-FEB-2000 (first entry)
XX
DE Human glycoprotein Ib/mouse IgG1Fc chimeric protein.
XX
KW Glycoprotein Ib; glycoallidin; detection; antithrombotic; binding;
KW von Willebrand factor; botrocetin; chimeric protein; immunoglobulin;
KW thrombotic disease; diagnosis; cardiac infarction; cerebral embolism.
XX
OS Homo sapiens.
OS Mus musculus.
OS Chimeric.
XX
XX Key Location/Qualifiers
XX Peptide 1..16 /label= signal
XX Protein 17..562 /label= human_glycoprotein_ib/mouse_IgG1Fc_chimeric
XX protein
XX
XX WO9954360-A1.
XX
XX 28-OCT-1999.
XX
XX 13-JAN-1999; 99WO-JP000089.
XX
XX 23-APR-1998; 98JP-00113962.
XX
XX (AJTN ) AJTNOMOTO CO INC.
XX
XX Fukuchi N, Futaki F, Kito M, Sato S, Kajiuira T, Ono Y, Ishii K;
XX Tanaka A, Shinozaki J, Jojima Y;
XX
XX WPI; 2000-013233/01.
XX
XX N-PSDB; AA235701.
XX
XX Direct quantitative detection of glycoallidin with immobilized von
XX Willebrand factor to bond with chimeric protein via inhibiting
XX glycoprotein Ib binding, for diagnosis of thrombotic diseases and
XX screening anti-thrombotic substances.
XX
XX Example 1; Page 70-71; 83pp; Japanese.
XX

```

XX A method has been developed for the detection of binding between the von
CC Willebrand factor and glycoprotein Ib or of the binding inhibition, in
CC which the von Willebrand factor immobilised in a reactor reacts with
CC glycoprotein Ib in the presence of a binding inducer to promote binding
CC between the von Willebrand factor and glycoprotein Ib. This method is for
CC the detection of glycoallidin as a means of thrombotic disease diagnosis
CC e.g. for cardiac infarction and cerebral embolism, and also for screening
CC substances with anti-thrombotic activity for the prevention and treatment
CC of thrombotic diseases. The method is direct, convenient and
CC quantitative, with reproducibility, and there is no need to construct a
CC monoclonal antibody for the assay. The present sequence is a human
CC glycoprotein Ib/mouse immunoglobulin gamma 1 Fc chimeric protein from the
CC present invention. (Updated on 12-SEP-2003 to standardise OS field)
XX SQ Sequence 562 AA;

Query Match 84.9%; Score 2411.5; DB 3; Length 562;
Best Local Similarity 79.5%; Pred. No. 1.1e-172;
Matches 449; Conservative 40; Mismatches 39; Indels 37; Gaps 4;
QY 1 MPLLALLLPSPLPHPHICVSVKASHLEVNCCKRNLTALPPDLPKDTTILHLENLLY 60
DB 1 MPLLALLLPSPLPHPHICVSVKASHLEVNCCKRNLTALPPDLPKDTTILHLENLLY 60
QY 61 TFSLATLMPYTRLTQNLDRCELTKLQVDGTLPLVLGTLDSLHNOQLSLPLGLQTLPALTV 120
DB 61 TFSLATLMPYTRLTQNLDRCELTKLQVDGTLPLVLGTLDSLHNOQLSLPLGLQTLPALTV 120
QY 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGELLTPPKLKLKSLANNLTLP 180
DB 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGELLTPPKLKLKSLANNLTLP 180
QY 181 AGLNGLENLDTLLQENSLYTIKPGFGSHLLPFAHGNPWLNCCEILYFRWLQDNA 240
DB 181 AGLNGLENLDTLLQENSLYTIKPGFGSHLLPFAHGNPWLNCCEILYFRWLQDNA 240
QY 241 ENVYVWKQVDVKAITSNVAQVQDSDKFPVYKPGKGCPTLGDGDTLDLYYPEDT 300
DB 241 ENVYVWKQVDVKAITSNVAQVQDSDKFPVYKPGKGCPTLGDGDTLDLYYPEDT 300
QY 301 EGDKVR-----PHT-----CPP--CPAPALGAPSVFL 326
DB 301 EGDKVRATRTVYKPTKAHTTTPWGLFYKSWSTASLDVPRDCGCKFCICTVPS--VSSVFI 357
QY 327 FPPKPKDTLMSRPEVTCVVVDYSHEDPEVKFNWYDGVVHNAKTPREEQYNSTRV 386
DB 358 FPPKPKDVTLLTPKVTCTVVVDISKDDPEVQFSWFDVDDVEVHTAQTPREEQNFSTRS 417
QY 387 VSVLTVLHQDWLNGKEYKCKVSNKALPVIIEKTIKSKAKGPQREPQVYTLPPSREEMTKNQ 446
DB 418 VSELPIMHQDWLNGKEYKCKVSNKALPVIIEKTIKSKAKGPQREPQVYTLPPSREEMTKNQ 446
QY 447 VSLCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFPLYSKLTVDKSRWQQGNV 506
DB 478 VSLTCLMTDFPDDITVEWQW 537
QY 507 FSCSWMEALHNYTKQSLSLSPGK 531
DB 538 FTCSVLHGLNHHHTKSLSPGK 562

RESULT 8
AA49935
ID AA49935 standard; protein; 568 AA.
XX
AC AA49935;
XX
DT 12-SEP-2003 (revised)
DT 01-FEB-2000 (first entry)
XX Human glycoprotein Ib/mouse IgG2aFc chimeric protein #2.
XX

KW Glycoprotein Ib; glycoallidin; detection; antithrombotic; binding;
KW Willebrand factor; botrocetin; chimeric protein; immunoglobulin;
KW thrombotic disease; diagnosis; cardiac infarction; cerebral embolism.
XX Homo sapiens.
OS Mus musculus.
OS Chimeric.
XX Key Location/Qualifiers
FT Peptide 1..16
FT Protein /label= signal
FT 17..568
FT /label= Human glycoprotein Ib/mouse_IgG2aFc_chimeric
FT protein
XX WO9954360-A1.
XX 28-OCT-1999.
XX 13-JAN-1999; 99WO-JP000089.
XX 23-APR-1998; 98JP-00113962.
XX (AJIN) AJINOMOTO CO INC.
XX Fukuchi N, Futaki F, Kito M, Sato S, Kajiura T, Ono Y, Ishii K;
XX Tanaka A, Shinozaki J, Jojima Y;
XX WPI; 2000-013233/01.
XX N-PSDB; AAZ35706.
XX Direct quantitative detection of glycoallidin with immobilized von
XX Willebrand factor to bond with chimeric protein via inhibiting
XX glycoprotein Ib binding, for diagnosis of thrombotic diseases and
XX screening anti-thrombotic substances.

Example 1; Page 77-79; 83pp; Japanese.
A method has been developed for the detection of binding between the von
Willebrand factor and glycoprotein Ib or of the binding inhibition, in
which the von Willebrand factor immobilised in a reactor reacts with
glycoprotein Ib in the presence of a binding inducer to promote binding
between the von Willebrand factor and glycoprotein Ib. This method is for
the detection of glycoallidin as a means of thrombotic disease diagnosis
e.g. for cardiac infarction and cerebral embolism, and also for screening
substances with anti-thrombotic activity for the prevention and treatment
of thrombotic diseases. The method is direct, convenient and
quantitative, with reproducibility, and there is no need to construct a
monoclonal antibody for the assay. The present sequence represents a
human glycoprotein Ib/mouse immunoglobulin gamma 2a Fc chimeric protein
from the present invention. (Updated on 12-SEP-2003 to standardise OS
field)

SQ Sequence 568 AA;
Query Match 83.9%; Score 2381.5; DB 3; Length 568;
Best Local Similarity 79.4%; Pred. No. 2e-170;
Matches 451; Conservative 29; Mismatches 51; Indels 37; Gaps 3;

QY 1 MPLLALLLPSPLPHPHICVSVKASHLEVNCCKRNLTALPPDLPKDTTILHLENLLY 60
DB 1 MPLLALLLPSPLPHPHICVSVKASHLEVNCCKRNLTALPPDLPKDTTILHLENLLY 60
QY 61 TFSLATLMPYTRLTQNLDRCELTKLQVDGTLPLVLGTLDSLHNOQLSLPLGLQTLPALTV 120
DB 61 TFSLATLMPYTRLTQNLDRCELTKLQVDGTLPLVLGTLDSLHNOQLSLPLGLQTLPALTV 120
QY 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGELLTPPKLKLKSLANNLTLP 180
DB 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGELLTPPKLKLKSLANNLTLP 180
QY 181 AGLNGLENLDTLLQENSLYTIKPGFGSHLLPFAHGNPWLNCCEILYFRWLQDNA 240
XX

Db 181 AGLNGLENDTLILLQENSLYTIKPGFFGSHLLPFAFLHGNPWLNCNCEILYFRWLQDNA 240
 QY 241 ENVYWKQVVDVKAVTSNVASVQCDNSDKFPVYKYPKGCGPTLGDGDTLDLYDYYPEEDT 300
 Db 241 ENVYWKQVVDVKAVTSNVASVQCDNSDKFPVYKYPKGCGPTLGDGDTLDLYDYYPEEDT 300
 QY 301 EGDQVR-----PHT-----CP-CPAPEALGADS 323
 Db 301 EGDQVRATRVKVPKTAHTTPMGLFYSWSTASLDSBERGPTIKPCPCPCPAPNLLGGPS 360
 QY 324 VLFPPPKDITLMSRTPEVTCVVVDVSHEDDPVKFNWYVDGVEVHNATKPREBOYNST 383
 Db 361 VLFPPPKDITLMSLPIVTCVVVDVSEDDPDVQISFWVNVVETAQTQTHREDYNST 420
 QY 384 YRVSVLTVLHQLNGEKYKCVSNKALVPVIEKTSKAKGQRPQVYVTLPPSREEMT 443
 Db 421 LRVVSALPIQHQMWSGKPKCVNNKDLPAPIERTISKFGSVRAQVYVLPPEEEMT 480
 QY 444 KNOVSLTCLVKGYFSPVDIAVWEESNGQPENNYKTTPVLDSDGSFPLYSKLTVDKSRWQ 503
 Db 481 KKQVTLTCMVTDMPEDIVYEVWNTNGKTELNYKTEPVLDSGSGYFMYSKLRVEKKNWVE 540
 QY 504 GNVFSCSVMEALHNNHYTKSLSPGK 531
 Db 541 RNSYSCSVHGLEHNNHHTYKSFRTGK 568

RESULT 9
 AA012135
 ID AA012135 standard; protein; 626 AA.
 AC AA012135;
 XX
 DT 03-JAN-2002 (first entry)
 DE Human glycoprotein Ib (platelet) alpha (GP1BA).
 KW Human; haplotyping; glycoprotein Ib (platelet) alpha protein; GP1BA;
 KW Bernard-Soulier syndrome; platelet-type von Willebrand disease; HIV;
 KW Alzheimer's disease; human immunodeficiency virus; SNP;
 KW single nucleotide polymorphism; chromosome 17pter-p12.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 72 /note= "Arg at this position is replaced with His due to
 single nucleotide polymorphism (SNP)"
 FT Misc-difference 86 /note= "Leu at this position is replaced with Phe due to
 single nucleotide polymorphism (SNP)"
 FT Misc-difference 161 /note= "Thr at this position is replaced with Met due to
 single nucleotide polymorphism (SNP)"
 FT Misc-difference 592 /note= "Arg at this position is replaced with His due to
 single nucleotide polymorphism (SNP)"
 FT Misc-difference 624 /note= "His at this position is replaced with Arg due to
 single nucleotide polymorphism (SNP)"
 XX
 PN WO200175065-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 03-APR-2001; 2001WO-US010671.
 XX
 PR 03-APR-2000; 2000US-0194341P.
 XX
 PA (GENA-) GENAISSANCE PHARM INC.
 XX
 PI Bentivegna SC, Choi JY, Klieem SE, Koshy B, Parks KB;
 XX

DR WPI; 2001-626427/72.
 DR N-PSDB; AAD20681.
 PT New haplotypes of the glycoprotein Ib platelet alpha polypeptide gene are
 useful for diagnosis and drug discovery for treating Bernard Soulier
 syndrome, platelet-type von Willebrand disease, HIV and Alzheimer's
 disease.
 PT
 XX Claim 29; Fig 3; 66pp; English.
 XX
 CC The invention relates to methods for haplotyping glycoprotein Ib
 (platelet) alpha polypeptide (GP1BA) gene of an individual. The method
 involves determining if the individual has one of the GP1BA haplotypes or
 haplotype pairs. The methods of the invention are useful for disease
 diagnosis and in the discovery and development of drugs for treating
 diseases associated with GP1BA activity e.g. Bernard-Soulier syndrome,
 platelet-type von Willebrand disease, HIV and Alzheimer's disease. The
 present sequence is human GP1BA protein. GP1BA gene is located on
 chromosome 17pter-p12
 XX
 SQ Sequence 626 AA;
 Query Match 56.9%; Score 1615.5; DB 4; Length 626;
 Best Local Similarity 84.2%; Pred. No. 1e-112;
 Matches 319; Conservative 5; Mismatches 20; Indels 35; Gaps 5;
 QY 1 MPILLLLLLLPSPLPHPICEVSKVASHLEVNCDEKNTALPDLPKDTTILHSENLLY 60
 Db 1 MPILLLLLLLPSPLPHPICEVSKVASHLEVNCDEKNTALPDLPKDTTILHSENLLY 60
 QY 61 TFSLATIMPYTRITQNLDRCELTKLVQDGTLPVLGTLDSHNLQSLPLGTLTALTV 120
 Db 61 TFSLATIMPYTRITQNLDRCELTKLVQDGTLPVLGTLDSHNLQSLPLGTLTALTV 120
 QY 121 LDVSNRLTSLPLGALRGELQELVYKGNELKTLPPGLTPTPKLEKLSLANNLTLP 180
 Db 121 LDVSNRLTSLPLGALRGELQELVYKGNELKTLPPGLTPTPKLEKLSLANNLTLP 180
 QY 181 AGLNGLENDTLILLQENSLYTIKPGFFGSHLLPFAFLHGNPWLNCNCEILYFRWLQDNA 240
 Db 181 AGLNGLENDTLILLQENSLYTIKPGFFGSHLLPFAFLHGNPWLNCNCEILYFRWLQDNA 240
 QY 241 ENVYWKQVVDVKAVTSNVASVQCDNSDKFPVYKYPKGCGPTLGDGDTLDLYDYYPEEDT 300
 Db 241 ENVYWKQVVDVKAVTSNVASVQCDNSDKFPVYKYPKGCGPTLGDGDTLDLYDYYPEEDT 300
 QY 301 EGDQVR-----PHTCP-----PCPA---PEALGAPSVLEFPK-- 330
 Db 301 EGDQVRATRVKVPKTAHTTPMGLFYSWSTASLDSQMPSSLHTQESTKEQITTFPRWT 360
 QY 331 PKDTL-----MISRTPEVT 344
 Db 361 PNFTLHMSITFSKTPKST 379

RESULT 10
 ADF69094
 ID ADF69094 standard; protein; 626 AA.
 XX
 AC ADF69094;
 XX
 DT 12-FEB-2004 (first entry)
 XX
 DE Human MP53 protein sequence SEQ ID NO:64.
 KW p53 pathway modulating agent; MP53; p53 modulator; cytostatic;
 KW gene therapy; cancer; human.
 XX
 OS Homo sapiens.
 XX
 PN WO2003083047-A2.
 XX
 PD 09-OCT-2003.

```
XX 28-FEB-2003; 2003WO-US006025.
XX
XX 01-MAR-2002; 2002US-0361196P.
XX
XX (EXEL-) EXELIXIS INC.
XX
XX Belvin M, Francis-Lang H, Friedman L, Plowman GD, Heuer TS, Li D;
XX Funke RP;
XX
XX WPI; 2003-812540/76.
XX N-P5DB; ADF69150.
XX
XX Identifying a candidate p53 pathway modulating agent for treating e.g.,
XX cancer by contacting an assay system comprising a MP53 polypeptide or
XX nucleic acid with a test agent and detecting a test agent-biased
XX activity.
XX
XX Example; SEQ ID NO 64; 406pp; English.
XX
XX The present invention describes a method for identifying a candidate p53
XX pathway modulating agent, which comprises:(a) providing an assay system
XX comprising a MP53 (modulator of p53) polypeptide or nucleic acid or its
XX fragment or derivative; (b) contacting the assay system with a test agent
XX under conditions where the system provides a reference activity except in
XX the presence of the test agent; and (c) detecting a test agent-biased
XX activity, where a difference between the test agent-biased activity and
XX the reference activity identifies the test agent as a candidate p53
XX pathway modulating agent. Also described: (1) modulating the p53 pathway
XX of a cell; (2) modulating the p53 pathway in a mammalian cell; and (3)
XX diagnosing a disease in a patient. MP53 has cytostatic activity, and can
XX be used in gene therapy. The method is useful for identifying a candidate
XX p53 pathway modulating agent for preparing a composition for diagnosing
XX or treating e.g., cancer. The present sequence represents a human MP53
XX protein, which is used in the exemplification of the present invention.
XX
XX Sequence 626 AA;
XX
XX Query Match 56.9%; Score 1615.5; DB 7; Length 626;
XX Best Local Similarity 84.2%; Pred. No. 1e-112;
XX Matches 319; Conservative 5; Mismatches 20; Indels 35; Gaps 5;
XX
XX Qy 1 MLLLLLLLLPSPLPHPICEVSKVASHLEVNCCKRNLTALPPDLPKDPTTLHLSENILLY 60
XX Db 1 MLLLLLLLLPSPLPHPICEVSKVASHLEVNCCKRNLTALPPDLPKDPTTLHLSENILLY 60
XX
XX Qy 61 TFSLATLMPYTRLTQNLNDRCELTKLQVDGTLPLVLTGLDLSHNLQSLPLLGQTLPALTV 120
XX Db 61 TFSLATLMPYTRLTQNLNDRCELTKLQVDGTLPLVLTGLDLSHNLQSLPLLGQTLPALTV 120
XX
XX Qy 121 LDVFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTPL 180
XX Db 121 LDVFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTPL 180
XX
XX Qy 181 AGLNGLENLTLLQENSLYTIKGFPGSHLLPFAFLHGNPWLNCBILYFRFLWQDNA 240
XX Db 181 AGLNGLENLTLLQENSLYTIKGFPGSHLLPFAFLHGNPWLNCBILYFRFLWQDNA 240
XX
XX Qy 241 ENVYVWKQVDKAVTSNVASVQCDNSDKFPVYKPGKCPGLDGBGDTLDYDYPEEDT 300
XX Db 241 ENVYVWKQVDKAVTSNVASVQCDNSDKFPVYKPGKCPGLDGBGDTLDYDYPEEDT 300
XX
XX Qy 301 EGDKVR-----PHTCP-----PCPA---PEALGAPSVFLFPKP-- 330
XX Db 301 EGDKVRATRVVKEPTKATHTPWGLFYSWSTASLDSPSSLHTQESTKEQTTPPPRWT 360
XX
XX Qy 331 PKDTL-----MISRTPEVT 344
XX Db 361 ENFTLHMSITFSKTPKST 379
XX
XX RESULT 11
XX ADJ68624
```

```
ID ADJ68624 standard; protein; 626 AA.
AC
XX ADJ68624;
XX
XX DT 06-MAY-2004 (first entry)
XX
XX DE Human heat mitochondrial protein as a therapeutic target SeqID430.
XX
XX KW mitochondrial; human; screening assay; diabetes mellitus;
XX KW Huntington's disease; osteoarthritis; LHON;
XX KW Leber's hereditary optic neuropathy; LHON;
XX KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
XX KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
XX KW neuroprotective; nootropic; antidiabetic; anticonvulsant; antiarthritic;
XX KW osteopathic; ophthalmological; cytostatic.
XX
XX OS Homo sapiens.
XX
XX FN WO2003087768-A2.
XX
XX PD 23-OCT-2003.
XX
XX PF 04-APR-2003; 2003WO-US010870.
XX
XX PR 12-APR-2002; 2002US-0372843P.
XX PR 17-JUN-2002; 2002US-0389987P.
XX PR 20-SEP-2002; 2002US-0412418P.
XX
XX (MITO-) MITOKOR.
XX (BUCK-) BUCK INST AGE RES.
XX
XX Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GW;
XX Warnock DB;
XX WPI; 2003-845369/78.
XX
XX Identifying a mitochondrial target for drug screening assays and for
XX treating diseases associated with altered mitochondrial function,
XX comprises detecting a modified polypeptide in a sample and correlating
XX with the disease.
XX
XX Claim 1; SEQ ID NO 430; 180pp; English.
XX
XX This invention relates to novel mitochondrial targets that can be used
XX for therapeutic intervention in treating a disease associated with
XX altered mitochondrial function. Specifically, it refers to a method for
XX identifying proteins of the human heart mitochondrial proteome that are
XX useful for drug screening assays, as well as therapeutic targets. The
XX present invention describes a method for identifying such proteins that
XX can be used in the treatment of various diseases associated with altered
XX mitochondrial function including diabetes mellitus, Huntington's disease,
XX osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
XX encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
XX ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
XX compositions have neuroprotective, nootropic, antidiabetic,
XX anticonvulsant, antiarthritic, osteopathic, ophthalmological and
XX cytostatic activities. This polypeptide sequence is a human heart
XX mitochondrial protein of the invention.
XX
XX Sequence 626 AA;
XX
XX Query Match 56.9%; Score 1615.5; DB 7; Length 626;
XX Best Local Similarity 84.2%; Pred. No. 1e-112;
XX Matches 319; Conservative 5; Mismatches 20; Indels 35; Gaps 5;
XX
XX Qy 1 MLLLLLLLLPSPLPHPICEVSKVASHLEVNCCKRNLTALPPDLPKDPTTLHLSENILLY 60
XX Db 1 MLLLLLLLLPSPLPHPICEVSKVASHLEVNCCKRNLTALPPDLPKDPTTLHLSENILLY 60
XX
XX Qy 61 TFSLATLMPYTRLTQNLNDRCELTKLQVDGTLPLVLTGLDLSHNLQSLPLLGQTLPALTV 120
XX Db 61 TFSLATLMPYTRLTQNLNDRCELTKLQVDGTLPLVLTGLDLSHNLQSLPLLGQTLPALTV 120
```

QY 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTHIPGILLTPPKLEKLSLANNLTLP 180
 Db 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTHIPGILLTPPKLEKLSLANNLTLP 180
 QY 181 AGLINGLENLDTLLQENSLYTIIPKGFSGSHLLPFAFLHGNPWLNCCEILYFRRWLQDNA 240
 Db 181 AGLINGLENLDTLLQENSLYTIIPKGFSGSHLLPFAFLHGNPWLNCCEILYFRRWLQDNA 240
 QY 241 ENVYWKQVVDVKAMTSNVASVQCDNSDKFPVYKPGKCPITLGDGDTLDLYDYPEEDT 300
 Db 241 ENVYWKQVVDVKAMTSNVASVQCDNSDKFPVYKPGKCPITLGDGDTLDLYDYPEEDT 300
 QY 301 EGDQVYR-----PHTCP-----PCPA---PEALGAPSVFLPFPK-- 330
 Db 301 EGDQVYRATRVVVKFPTKAHTTPWGLFYSWSTASLDSQMPSSLHTQESTKEQTTFPPRWT 360
 QY 331 PKDTL-----MISRTPEVT 344
 Db 361 PNFTLHMSITFSKTPKST 379
 RESULT 12
 ID ADH40324
 XX ADH40324 standard; protein; 626 AA.
 AC ADH40324;
 XX
 DT 11-MAR-2004 (first entry)
 XX Human platelet glycoprotein IB alpha chain precursor CD42b.
 XX human; platelet glycoprotein IB alpha chain precursor; CD42b; cytostatic;
 KW vaccine; SNP profile; cancer; leukaemia.
 XX Homo sapiens.
 XX WO2003106692-A2.
 PD 24-DEC-2003.
 PF 13-JUN-2003; 2003WO-EP006251.
 XX 13-JUN-2002; 2002EP-00013423.
 PR (MERE) MERCK PATENT GMBH.
 PA Strittmatter W, Moll H;
 XX
 PI
 XX WPI; 2004-082200/08.
 DR
 XX Providing allelic variant epitope of protein based on single nucleotide
 PT polymorphism by defining target protein, screening database of protein,
 PT identifying, selecting allelic variant protein, creating variant
 PT epitopes.
 XX
 PS Example 4; Page 67; 119pp; English.
 XX
 CC The invention relates to a novel method for providing epitopes of allelic
 CC variants of antigenic proteins from specific species based on single
 CC nucleotide polymorphism (SNP), by defining target protein/peptide or its
 CC subset, screening database of DNA encoding target protein, identifying,
 CC selecting allelic peptide/protein variants, expression product or its
 CC fragment encoded by DNA sequence having SNP, creating variant epitopes,
 CC selecting epitopes binding to MHC protein. A protein of the invention has
 CC cytostatic activity, and may have a use in a vaccine. The method is
 CC useful for generating a SNP profile of one or more individuals from a
 CC given species by applying the method for several protein from the
 CC individuals, where the SNP profile was related to disease, preferably
 CC cancer. This is useful for diagnosing a disease in an individual by
 CC generating the SNP-related polymorphic profile. A method of the invention
 CC is useful for transplanting haematopoietic stem cells from a donor to a
 CC recipient and treating cancer, preferably leukaemia, and for determining
 CC the progression, regression or onset of a treated disease. The present

CC sequence is used in the exemplification of the invention.
 XX
 SQ Sequence 626 AA;
 Query Match 56.9%; Score 1615.5; DB 8; Length 626;
 Best Local Similarity 84.2%; Pred. No. 18-112;
 Matches 319; Conservative 5; Mismatches 20; Indels 35; Gaps 5;
 QY 1 MPILLALLLLSPLPHPICEVSKVASHLEVNCDEKNTALPPDLPKDTTILHLSNLLY 60
 Db 1 MPILLALLLLSPLPHPICEVSKVASHLEVNCDEKNTALPPDLPKDTTILHLSNLLY 60
 QY 61 TFSLATMPTYTRLTQNLDRCELTKLQVDTLPVLGTLDSHNQLQSLPLGOTLPALT 120
 Db 61 TFSLATMPTYTRLTQNLDRCELTKLQVDTLPVLGTLDSHNQLQSLPLGOTLPALT 120
 QY 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTHIPGILLTPPKLEKLSLANNLTLP 180
 Db 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTHIPGILLTPPKLEKLSLANNLTLP 180
 QY 181 AGLINGLENLDTLLQENSLYTIIPKGFSGSHLLPFAFLHGNPWLNCCEILYFRRWLQDNA 240
 Db 181 AGLINGLENLDTLLQENSLYTIIPKGFSGSHLLPFAFLHGNPWLNCCEILYFRRWLQDNA 240
 QY 241 ENVYWKQVVDVKAMTSNVASVQCDNSDKFPVYKPGKCPITLGDGDTLDLYDYPEEDT 300
 Db 241 ENVYWKQVVDVKAMTSNVASVQCDNSDKFPVYKPGKCPITLGDGDTLDLYDYPEEDT 300
 QY 301 EGDQVYR-----PHTCP-----PCPA---PEALGAPSVFLPFPK-- 330
 Db 301 EGDQVYRATRVVVKFPTKAHTTPWGLFYSWSTASLDSQMPSSLHTQESTKEQTTFPPRWT 360
 QY 331 PKDTL-----MISRTPEVT 344
 Db 361 PNFTLHMSITFSKTPKST 379
 RESULT 13
 ID ADQ39833
 XX ADQ39833 standard; protein; 626 AA.
 AC ADQ39833;
 XX
 DT 18-NOV-2004 (first entry)
 XX Human myocardial infarction-associated gene derived protein, SEQ ID 1496.
 DE Myocardial infarction; detection; single nucleotide polymorphism; SNP;
 KW cardiant; gene therapy; human.
 XX Homo sapiens.
 XX WO2004058052-A2.
 PD 15-JUL-2004.
 XX
 PF 22-DEC-2003; 2003WO-US040978.
 PR 20-DEC-2002; 2002US-0434778P.
 PR 10-MAR-2003; 2003US-0453135P.
 PR 30-APR-2003; 2003US-0466412P.
 PR 23-SEP-2003; 2003US-0504955P.
 XX (APPL-) APPLERA CORP.
 PA Cargill M, Devlin J, Iakoubova O;
 PI WPI; 2004-533949/51.
 XX N-PSDB; ADQ39005.
 DR Identifying an individual who has an altered risk for developing
 PT myocardial infarction by detecting a single nucleotide polymorphism in
 PT the individual's nucleic acids.

XX PS Claim 10; SEQ ID NO 1496; 145pp; English.

XX CC The invention relates to a novel method for identifying an individual who has an altered risk for developing myocardial infarction. The method comprises detecting a single nucleotide polymorphism (SNP) in any one of the nucleotide sequences given in the specification in the individual's nucleic acids, where the presence of the SNP is correlated with an altered risk for myocardial infarction in the individual. The invention further comprises: an isolated nucleic acid molecule comprising at least 8 contiguous nucleotides where one of the nucleotides is an SNP given in the specification or its complement and encoding any one of the amino acid sequences given in the specification; an isolated polypeptide comprising an amino acid sequence given in the specification; an antibody that specifically binds to the polypeptide or its antigen-binding fragment; an amplified polynucleotide containing an SNP given in the specification and which is between about 16 and 1000 nucleotides in length; a kit for detecting an SNP in a nucleic acid, comprising the polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a nucleic acid molecule; a method of detecting a variant polypeptide; and a method for identifying an agent useful in treating or preventing myocardial infarction. The novel detection method has cardiac activity. The nucleic acids of the invention may be used in gene therapy. The method is useful in identifying an individual who has an increased or decreased risk for developing myocardial infarction and for preparing a composition for treating or preventing myocardial infarction. This sequence represents the protein of a human myocardial infarction-associated gene containing one or more SNPs of the invention. Note: This sequence was not shown in the specification. The sequence has come from an electronic sequence listing downloaded from the WIPO website.

XX SQ Sequence 626 AA;

Query Match 56.9%; Score 1615.5; DB 8; Length 626;
Best Local Similarity 84.2%; Pred. No. 1e-112;
Matches 319; Conservative 5; Mismatches 20; Indels 35; Gaps 5;

Qy 1 MPLLILLPLSPHPHPICEVSKVASHLEVNCCKRNLTALPPDLPKDTHLHSENL 60
Db 1 MPLLILLPLSPHPHPICEVSKVASHLEVNCCKRNLTALPPDLPKDTHLHSENL 60

Qy 61 TFSLATLMPYTRLTQNLDRCELTKLQVDGTLPLVGLTDLSHNQLQSLPLLGQTLPALTV 120
Db 61 TFSLATLMPYTRLTQNLDRCELTKLQVDGTLPLVGLTDLSHNQLQSLPLLGQTLPALTV 120

Qy 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTPL 180
Db 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTPL 180

Qy 181 AGLNGLENLDTLLQENSLYTI PKGFFGSHLLPPAFHGNPNWLCNCEILYFRWLQDNA 240
Db 181 AGLNGLENLDTLLQENSLYTI PKGFFGSHLLPPAFHGNPNWLCNCEILYFRWLQDNA 240

Qy 241 ENVYVKQVDVKAVTSNVASVQCDNSDKFPVYKPGKCPGLTGEGDGLDYYPEDT 300
Db 241 ENVYVKQVDVKAVTSNVASVQCDNSDKFPVYKPGKCPGLTGEGDGLDYYPEDT 300

Qy 301 EGDKVR-----PHTCF-----PCPA---PEALGAPSVFLFPKK-- 330
Db 301 EGDKVRATRVVVKFPTKAHTTPWGLFYSWSTASLDQSPSSLHPQTQSTKEQTTPFPKRW 360

Qy 331 PKDTL-----MISRTPEVT 344
Db 361 PNFTLHMSITFSKTPKST 379

RESULT 14
ABB78244
ID ABB78244 standard; protein; 290 AA.
XX AC ABB78244;
XX DT 25-NOV-2002 (first entry)

XX DE Amino acid sequence of glycoprotein 1B-alpha polypeptide GPIb290/2V.
XX KW Glycoprotein 1B-alpha; GPIb; immunoglobulin; Ig; platelet adherence;
KW leukocyte; platelet activation; ischaemic heart disease;
KW acute myocardial infarction; stroke; venous thrombosis; atherosclerosis;
KW arterial thrombosis; angina; vascular condition; vascular inflammation;
KW thrombosis; angioplasty; restenosis.
XX OS Unidentified.
XX PN WO200263003-A2.
XX PD 15-AUG-2002.
XX PF 06-FEB-2002; 2002WO-US003549.
XX PR 06-FEB-2001; 2001US-0266838P.
XX PA (GEMY) GENETICS INST LLC.
XX PI Shaw GD, Kumar R, Sako DS, Mcdonagh T, Sullivan FX;
XX WPI; 2002-657537/70.
XX DR New glycoprotein 1B alpha fusion polypeptides, useful for treating a disorder associated with platelet activation e.g. ischemic heart disease, stroke, venous or arterial thrombosis or atherosclerosis.
XX PT
XX PT
XX PS Disclosure; Page 11; 45pp; English.
XX CC ABB78240-45 represent glycoprotein 1B-alpha (GPIb) polypeptides, which are used to produce fusion proteins with an immunoglobulin (Ig) polypeptide. The fusion proteins inhibit the adherence of platelets to leukocytes. The fusion polypeptides are useful for treating a disorder associated with platelet activation e.g. ischaemic heart disease, acute myocardial infarction, stroke, venous thrombosis, atherosclerosis, arterial thrombosis or unstable angina. They can also be used to treat vascular conditions associated with vascular inflammation, thrombosis, and angioplasty-related restenosis
XX SQ Sequence 290 AA;

Query Match 54.1%; Score 1536; DB 5; Length 290;
Best Local Similarity 100.0%; Pred. No. 3.6e-107;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 HPICEVSKVASHLEVNCCKRNLTALPPDLPKDTHLHSENLTYTFSLATLMPYTRLTQ 76
Db 1 HPICEVSKVASHLEVNCCKRNLTALPPDLPKDTHLHSENLTYTFSLATLMPYTRLTQ 60

Qy 77 NLDRCELTKLQVDGTLPLVGLTDLSHNQLQSLPLLGQTLPALTVLDVSFNRLTSLPLGAL 136
Db 61 NLDRCELTKLQVDGTLPLVGLTDLSHNQLQSLPLLGQTLPALTVLDVSFNRLTSLPLGAL 120

Qy 137 RGLGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTPLPAGLNGLENLDTLLQ 196
Db 121 RGLGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTPLPAGLNGLENLDTLLQ 180

Qy 197 ENSLYTTPKGFPGSHLLPPAFHGNPNWLCNCEILYFRWLQDNAENVYVKQVDVKAVT 256
Db 181 ENSLYTTPKGFPGSHLLPPAFHGNPNWLCNCEILYFRWLQDNAENVYVKQVDVKAVT 240

Qy 257 SNVASVQCDNSDKFPVYKPGKCPGLTGEGDGLDYYPEDTEGDKVR 306
Db 241 SNVASVQCDNSDKFPVYKPGKCPGLTGEGDGLDYYPEDTEGDKVR 290

RESULT 15
ABB78240
ID ABB78240 standard; protein; 301 AA.
XX AC ABB78240;

Search completed: May 24, 2005, 06:00:09
Job time : 167 secs

```

XX 25-NOV-2002 (first entry)
XX DE
XX Amino acid sequence of glycoprotein 1B-alpha polypeptide GP1b302.
XX
XX Glycoprotein 1B-alpha; GPIb; immunoglobulin; Ig; platelet adherence;
XX KW leukocyte; platelet activation; ischaemic heart disease;
XX KW acute myocardial infarction; stroke; venous thrombosis; atherosclerosis;
XX KW arterial thrombosis; angina; vascular condition; vascular inflammation;
XX KW thrombosis; angioplasty; restenosis.
XX OS
XX Unidentified.
XX PN WO200263003-A2.
XX PD
XX 15-AUG-2002.
XX
XX 06-FEB-2002; 2002WO-US003549.
XX PF
XX
XX 06-FEB-2001; 2001US-0266838P.
XX PR
XX (GEMY ) GENETICS INST LLC.
XX PA
XX
XX Shaw GD, Kumar R, Sako DS, Mcdonagh T, Sullivan FX;
XX PI
XX WPI; 2002-657537/70.
XX DR
XX
XX New glycoprotein 1b alpha fusion polypeptides, useful for treating a
XX PT disorder associated with platelet activation e.g. ischemic heart disease,
XX PT stroke, venous or arterial thrombosis or atherosclerosis.
XX PT
XX
XX Disclosure; Page 10; 45pp; English.
XX PS
XX
XX ABB78240-45 represent glycoprotein 1B-alpha (GPIb) polypeptides, which
XX CC are used to produce fusion proteins with an immunoglobulin (Ig)
XX CC polypeptide. The fusion proteins inhibit the adherence of platelets to
XX CC leukocytes. The fusion polypeptides are useful for treating a disorder
XX CC associated with platelet activation e.g. ischaemic heart disease, acute
XX CC myocardial infarction, stroke, venous thrombosis, atherosclerosis,
XX CC arterial thrombosis or unstable angina. They can also be used to treat
XX CC vascular conditions associated with vascular inflammation, thrombosis,
XX CC and angioplasty-related restenosis
XX CC
XX
XX Sequence 301 AA;
XX
Query Match 53.8%; Score 1527; DB 5; Length 301;
Best Local Similarity 93.0%; Pred. No. 1.8e-106;
Matches 292; Conservative 2; Mismatches 6; Indels 14; Gaps 1;
QY 17 HPICEVSKVASHLEVNCDEKNTALPPDLPKDTTILHLSNLYTFSLATLMEYTRLTQL 76
Db 1 HPICEVSKVASHLEVNCDEKNTALPPDLPKDTTILHLSNLYTFSLATLMEYTRLTQL 60
QY 77 NLDRCBLTKLVQDGTLPVLGTLDSHNQSLPLGQTLPALTVLDVSNFRLTSLPLGAL 136
Db 61 NLDRCBLTKLVQDGTLPVLGTLDSHNQSLPLGQTLPALTVLDVSNFRLTSLPLGAL 120
QY 137 RGIGELQELYLKGNELKLTLPGLTTPKLEKLSLANNLTLPAGLNGLENLDTLLIQ 196
Db 121 RGIGELQELYLKGNELKLTLPGLTTPKLEKLSLANNLTLPAGLNGLENLDTLLIQ 180
QY 197 ENSLYTIPKGFPGSHLLPFAFLHGNFWLNCCEILYFRRWLQDNAENVYVKQVVDKAVT 256
Db 181 ENSLYTIPKGFPGSHLLPFAFLHGNFWLNCCEILYFRRWLQDNAENVYVKQVVDKAVT 240
QY 257 SNVAVQCDSNDKFPVYKPGKCPYTLGDEGDTLDYDYYPEDETEGDKVRPHPTCPPCPAP 316
Db 241 SNVAVQCDSNDKFPVYKPGKCPYTLGDEGDTLDYDYYPEDETEGDKVR----- 290
QY 317 EALGAPSVFLFPPK 330
Db 291 ----ATRTWKFPK 300
```

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2005, 06:00:18 ; Search time 136 Seconds
(without alignments)
1306.055 Million cell updates/sec

Title: US-10-068-426-5

Perfect score: 2839

Sequence: 1 MFLLLLLLPPLPHPTIC.....MHEALNNHYTKSLSPQK 531

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1434725 seqs, 334507595 residues

Total number of hits satisfying chosen parameters: 1434725

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2839	100.0	531	14	US-10-068-426-5 Sequence 5, Appli
2	2839	100.0	531	15	US-10-382-758-5 Sequence 5, Appli
3	2829	99.6	531	14	US-10-068-426-4 Sequence 4, Appli
4	2829	99.6	531	15	US-10-382-758-4 Sequence 4, Appli
5	2823	99.4	531	14	US-10-068-426-6 Sequence 6, Appli
6	2823	99.4	531	15	US-10-382-758-6 Sequence 6, Appli
7	2822.5	99.4	544	14	US-10-068-426-3 Sequence 3, Appli
8	2822.5	99.4	544	15	US-10-382-758-3 Sequence 3, Appli
9	2812.5	99.1	544	14	US-10-068-426-1 Sequence 1, Appli
10	2812.5	99.1	544	15	US-10-382-758-1 Sequence 1, Appli
11	2812.5	99.1	544	14	US-10-068-426-2 Sequence 2, Appli
12	2812.5	99.1	544	15	US-10-382-758-2 Sequence 2, Appli
13	1615.5	56.9	626	16	US-10-408-765A-430 Sequence 430, App

14	1615.5	56.9	626	17	US-10-741-600-1496	Sequence 1496, Ap
15	1536	54.1	290	14	US-10-068-426-11	Sequence 11, Appl
16	1536	54.1	290	15	US-10-382-758-11	Sequence 11, Appl
17	1536	54.1	290	17	US-10-868-371-2	Sequence 2, Appli
18	1528	53.8	307	17	US-10-868-371-7	Sequence 7, Appli
19	1526.5	53.8	302	14	US-10-068-426-7	Sequence 7, Appli
20	1526.5	53.8	302	15	US-10-382-758-7	Sequence 7, Appli
21	1526.5	53.8	302	17	US-10-868-371-3	Sequence 3, Appli
22	1526	53.8	290	14	US-10-068-426-10	Sequence 10, Appl
23	1526	53.8	290	15	US-10-382-758-10	Sequence 10, Appl
24	1526	53.8	290	17	US-10-868-371-1	Sequence 1, Appli
25	1521.5	53.6	302	14	US-10-068-426-8	Sequence 8, Appli
26	1521.5	53.6	302	15	US-10-382-758-8	Sequence 8, Appli
27	1521.5	53.6	302	17	US-10-868-371-4	Sequence 4, Appli
28	1520	53.5	290	14	US-10-068-426-12	Sequence 12, Appl
29	1520	53.5	290	15	US-10-382-758-12	Sequence 12, Appl
30	1520	53.5	290	17	US-10-868-371-6	Sequence 6, Appli
31	1514	53.3	301	14	US-10-068-426-9	Sequence 9, Appli
32	1514	53.3	301	15	US-10-382-758-9	Sequence 9, Appli
33	1514	53.3	301	17	US-10-868-371-5	Sequence 5, Appli
34	1279	45.1	800	14	US-10-193-477-220	Sequence 220, App
35	1237	43.6	313	10	US-09-935-144-36	Sequence 36, Appl
36	1237	43.6	313	10	US-09-825-580-4	Sequence 4, Appli
37	1237	43.6	313	14	US-10-211-786-4	Sequence 4, Appli
38	1232.5	43.4	470	14	US-10-264-634-33	Sequence 33, Appl
39	1230.5	43.3	401	9	US-09-859-361-9	Sequence 9, Appli
40	1230	43.3	388	9	US-09-784-623-16	Sequence 16, Appl
41	1230	43.3	449	14	US-10-323-268-23	Sequence 23, Appl
42	1230	43.3	502	14	US-10-363-427-24	Sequence 24, Appl
43	1229.5	43.3	542	15	US-10-343-063A-25	Sequence 25, Appl
44	1228.5	43.3	430	17	US-10-841-250-22	Sequence 22, Appl
45	1228.5	43.3	528	15	US-10-431-359-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1

US-10-068-426-5
; Sequence 5, Application US/10068426
; Publication No. US20030091576A1
; GENERAL INFORMATION:
; APPLICANT: Shaw, Gray D.
; APPLICANT: Sako, Dianne S.
; APPLICANT: Kumar, Ravindra
; APPLICANT: Sullivan, Francis
; APPLICANT: McDonagh, Tom
; TITLE OF INVENTION: Platelet Glycoprotein IB Alpha Fusion Polypeptides and
; FILE REFERENCE: Methods of Use Thereof
; CURRENT APPLICATION NUMBER: US/10/068,426
; CURRENT FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/266,838
; PRIOR FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 5
; LENGTH: 531
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(531)
; OTHER INFORMATION: GPIIb290/2V-Ig
US-10-068-426-5

Query Match 100.0%; Score 2839; DB 14; Length 531;
Best Local Similarity 100.0%; Pred. NO. 8e-201;
Matches 531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFLLLLLLPPLPHPTICVSKVASHLEVNCNKNTALPPDLPKDTTILHSENLLY 60
DB 1 MFLLLLLLPPLPHPTICVSKVASHLEVNCNKNTALPPDLPKDTTILHSENLLY 60

```
Qy 61 TFSLATLMPYTRLTQLNDRCELTKLQVDGTLPLVGLTDLDSHNQLSLPLLAGQTLPALTV 120
Db 61 TFSLATLMPYTRLTQLNDRCELTKLQVDGTLPLVGLTDLDSHNQLSLPLLAGQTLPALTV 120
Qy 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLP 180
Db 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLP 180
Qy 181 AGLNGLENLDTLLQENSLYTIKPGFFGSHLLPFAFLHGNPWLNCNCELILYFRWLQDNA 240
Db 181 AGLNGLENLDTLLQENSLYTIKPGFFGSHLLPFAFLHGNPWLNCNCELILYFRWLQDNA 240
Qy 241 ENVVWVKQVVDVKAITSNVASVQCDNSDKFPVYKPGKGCPTLGDGDTLDLYYPEEDT 300
Db 241 ENVVWVKQVVDVKAITSNVASVQCDNSDKFPVYKPGKGCPTLGDGDTLDLYYPEEDT 300
Qy 301 EGDKVRPHTCPCPAPEALGAPSVFLPPPKPDTLMISRTPEVTCVVVDVSHEDPEVKFN 360
Db 301 EGDKVRPHTCPCPAPEALGAPSVFLPPPKPDTLMISRTPEVTCVVVDVSHEDPEVKFN 360
Qy 361 WYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPVPKEITI 420
Db 361 WYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPVPKEITI 420
Qy 421 SKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPP 480
Db 421 SKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPP 480
Qy 481 VLDSGGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNNHYTKQSLSLSPGK 531
Db 481 VLDSGGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNNHYTKQSLSLSPGK 531
```

RESULT 2

```
US-10-382-758-5
; Sequence 5, Application US/10382758
; Publication No. US20030232047A1
; GENERAL INFORMATION:
; APPLICANT: Shaw, Gray D.
; APPLICANT: Sako, Dianne S.
; APPLICANT: Kumar, Ravindra
; APPLICANT: Sullivan, Francis
; APPLICANT: McDonagh, Tom
; TITLE OF INVENTION: Platelet Glycoprotein IB Alpha Fusion Polypeptides and
; FILE REFERENCE: 22058-503
; CURRENT APPLICATION NUMBER: US/10/382,758
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: US/10/068,426
; PRIOR FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/266,838
; PRIOR FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 5
; LENGTH: 531
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(531)
; OTHER INFORMATION: GP1b290/2V-Ig
US-10-382-758-5
```

```
Query Match 100.0%; Score 2839; DB 15; Length 531;
Best Local Similarity 100.0%; Pred. No. 8e-201;
Matches 531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MLLLLLLLLPSPLPHPICEVSKVASHLVNCDKRNLTALPPDLPKDTTILHSENLLY 60
Db 1 MLLLLLLLLPSPLPHPICEVSKVASHLVNCDKRNLTALPPDLPKDTTILHSENLLY 60
```

```
Qy 61 TFSLATLMPYTRLTQLNDRCELTKLQVDGTLPLVGLTDLDSHNQLSLPLLAGQTLPALTV 120
Db 61 TFSLATLMPYTRLTQLNDRCELTKLQVDGTLPLVGLTDLDSHNQLSLPLLAGQTLPALTV 120
Qy 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLP 180
Db 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLP 180
Qy 181 AGLNGLENLDTLLQENSLYTIKPGFFGSHLLPFAFLHGNPWLNCNCELILYFRWLQDNA 240
Db 181 AGLNGLENLDTLLQENSLYTIKPGFFGSHLLPFAFLHGNPWLNCNCELILYFRWLQDNA 240
Qy 241 ENVVWVKQVVDVKAITSNVASVQCDNSDKFPVYKPGKGCPTLGDGDTLDLYYPEEDT 300
Db 241 ENVVWVKQVVDVKAITSNVASVQCDNSDKFPVYKPGKGCPTLGDGDTLDLYYPEEDT 300
Qy 301 EGDKVRPHTCPCPAPEALGAPSVFLPPPKPDTLMISRTPEVTCVVVDVSHEDPEVKFN 360
Db 301 EGDKVRPHTCPCPAPEALGAPSVFLPPPKPDTLMISRTPEVTCVVVDVSHEDPEVKFN 360
Qy 361 WYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPVPKEITI 420
Db 361 WYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPVPKEITI 420
Qy 421 SKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPP 480
Db 421 SKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPP 480
Qy 481 VLDSGGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNNHYTKQSLSLSPGK 531
Db 481 VLDSGGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNNHYTKQSLSLSPGK 531
```

RESULT 3

```
US-10-068-426-4
; Sequence 4, Application US/10068426
; Publication No. US20030091576A1
; GENERAL INFORMATION:
; APPLICANT: Shaw, Gray D.
; APPLICANT: Sako, Dianne S.
; APPLICANT: Kumar, Ravindra
; APPLICANT: Sullivan, Francis
; APPLICANT: McDonagh, Tom
; TITLE OF INVENTION: Platelet Glycoprotein IB Alpha Fusion Polypeptides and
; FILE REFERENCE: 22058-503
; CURRENT APPLICATION NUMBER: US/10/068,426
; CURRENT FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/266,838
; PRIOR FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4
; LENGTH: 531
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(531)
; OTHER INFORMATION: GP1b290-Ig
US-10-068-426-4
```

```
Query Match 99.6%; Score 2829; DB 14; Length 531;
Best Local Similarity 99.6%; Pred. No. 4.4e-200;
Matches 529; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MLLLLLLLLPSPLPHPICEVSKVASHLVNCDKRNLTALPPDLPKDTTILHSENLLY 60
Db 1 MLLLLLLLLPSPLPHPICEVSKVASHLVNCDKRNLTALPPDLPKDTTILHSENLLY 60
Qy 61 TFSLATLMPYTRLTQLNDRCELTKLQVDGTLPLVGLTDLDSHNQLSLPLLAGQTLPALTV 120
Db 61 TFSLATLMPYTRLTQLNDRCELTKLQVDGTLPLVGLTDLDSHNQLSLPLLAGQTLPALTV 120
```



```
QY 121 LDVSFNRLLTSLPLGALRGIGLQELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTTELP 180
DB 121 LDVSFNRLLTSLPLGALRGIGLQELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTTELP 180
QY 181 AGLLNGLENLDTLQLQENSLLYTI PKGFFGSHLLPFAFLHGNPWLNCCEILYFRRWLQDNA 240
DB 181 AGLLNGLENLDTLQLQENSLLYTI PKGFFGSHLLPFAFLHGNPWLNCCEILYFRRWLQDNA 240
QY 241 ENYYVWKQVVDVKAITSNVASVQCDNSDKFPVYKYPKGKCPITLGDGDTLDLYDYYPEEDT 300
DB 241 ENYYVWKQVVDVKAITSNVASVQCDNSDKFPVYKYPKGKCPITLGDGDTLDLYDYYPEEDT 300
QY 301 EGDKVRPHTCPPCPAPALGAPSVFLFPKPKDITLMISRTPEVTCVVVDVSHEDPEVKEN 360
DB 301 EGDKVRPHTCPPCPAPALGAPSVFLFPKPKDITLMISRTPEVTCVVVDVSHEDPEVKEN 360
QY 361 WYVDGVEVHNATKPREEQNSTYRVSVLTVLHQDLNKGKEYCKVSNKALPVPLEKTI 420
DB 361 WYVDGVEVHNATKPREEQNSTYRVSVLTVLHQDLNKGKEYCKVSNKALPVPLEKTI 420
QY 421 SKAKGPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTP 480
DB 421 SKAKGPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTP 480
QY 481 VLDSGSPFLYSLKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 531
DB 481 VLDSGSPFLYSLKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 531
```

RESULT 4

```
US-10-382-758-4
; Sequence 4, Application US/10382758
; Publication No. US20030232047A1
; GENERAL INFORMATION:
; APPLICANT: Shaw, Gray D.
; APPLICANT: Sako, Dianne S.
; APPLICANT: Kumar, Ravindra
; APPLICANT: Sullivan, Francis
; APPLICANT: McDonagh, Tom
; TITLE OF INVENTION: Platelet Glycoprotein IB Alpha Fusion Polypeptides and
; TITLE OF INVENTION: Methods of Use Thereof
; FILE REFERENCE: 22058-503
; CURRENT APPLICATION NUMBER: US/10/382,758
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: US/10/068,426
; PRIOR FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/266,838
; PRIOR FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4
; LENGTH: 531
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(531)
; OTHER INFORMATION: GPIb290-Ig
US-10-382-758-4
```

```
Query Match          99.6%; Score 2829; DB 15; Length 531;
Best Local Similarity 99.6%; Pred. No. 4.4e-200;
Matches 529; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 MPLLALLLPSPLPHPICEVSKVASHLEVNCCKENLTALPPDLPKDTTILHLSNLLY 60
DB 1 MPLLALLLPSPLPHPICEVSKVASHLEVNCCKENLTALPPDLPKDTTILHLSNLLY 60
QY 61 TFSLATIMPYTRLTQNLDRCELTKLQVDGTLPLVGLTDLSHNQSLPLLGOTLPALT 120
DB 61 TFSLATIMPYTRLTQNLDRCELTKLQVDGTLPLVGLTDLSHNQSLPLLGOTLPALT 120
```

```
QY 121 LDVSFNRLLTSLPLGALRGIGLQELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTTELP 180
DB 121 LDVSFNRLLTSLPLGALRGIGLQELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTTELP 180
QY 181 AGLLNGLENLDTLQLQENSLLYTI PKGFFGSHLLPFAFLHGNPWLNCCEILYFRRWLQDNA 240
DB 181 AGLLNGLENLDTLQLQENSLLYTI PKGFFGSHLLPFAFLHGNPWLNCCEILYFRRWLQDNA 240
QY 241 ENYYVWKQVVDVKAITSNVASVQCDNSDKFPVYKYPKGKCPITLGDGDTLDLYDYYPEEDT 300
DB 241 ENYYVWKQVVDVKAITSNVASVQCDNSDKFPVYKYPKGKCPITLGDGDTLDLYDYYPEEDT 300
QY 301 EGDKVRPHTCPPCPAPALGAPSVFLFPKPKDITLMISRTPEVTCVVVDVSHEDPEVKEN 360
DB 301 EGDKVRPHTCPPCPAPALGAPSVFLFPKPKDITLMISRTPEVTCVVVDVSHEDPEVKEN 360
QY 361 WYVDGVEVHNATKPREEQNSTYRVSVLTVLHQDLNKGKEYCKVSNKALPVPLEKTI 420
DB 361 WYVDGVEVHNATKPREEQNSTYRVSVLTVLHQDLNKGKEYCKVSNKALPVPLEKTI 420
QY 421 SKAKGPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTP 480
DB 421 SKAKGPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTP 480
QY 481 VLDSGSPFLYSLKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 531
DB 481 VLDSGSPFLYSLKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 531
```

RESULT 5

```
US-10-068-426-6
; Sequence 6, Application US/10068426
; Publication No. US20030091576A1
; GENERAL INFORMATION:
; APPLICANT: Shaw, Gray D.
; APPLICANT: Sako, Dianne S.
; APPLICANT: Kumar, Ravindra
; APPLICANT: Sullivan, Francis
; APPLICANT: McDonagh, Tom
; TITLE OF INVENTION: Platelet Glycoprotein IB Alpha Fusion Polypeptides and
; TITLE OF INVENTION: Methods of Use Thereof
; FILE REFERENCE: 22058-503
; CURRENT APPLICATION NUMBER: US/10/068,426
; CURRENT FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/266,838
; PRIOR FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 6
; LENGTH: 531
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(531)
; OTHER INFORMATION: GPIb290/1A-Ig
US-10-068-426-6
```

```
Query Match          99.4%; Score 2823; DB 14; Length 531;
Best Local Similarity 99.4%; Pred. No. 1.2e-199;
Matches 528; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 MPLLALLLPSPLPHPICEVSKVASHLEVNCCKENLTALPPDLPKDTTILHLSNLLY 60
DB 1 MPLLALLLPSPLPHPICEVSKVASHLEVNCCKENLTALPPDLPKDTTILHLSNLLY 60
QY 61 TFSLATIMPYTRLTQNLDRCELTKLQVDGTLPLVGLTDLSHNQSLPLLGOTLPALT 120
DB 61 TFSLATIMPYTRLTQNLDRCELTKLQVDGTLPLVGLTDLSHNQSLPLLGOTLPALT 120
QY 121 LDVSFNRLLTSLPLGALRGIGLQELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTTELP 180
DB 121 LDVSFNRLLTSLPLGALRGIGLQELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTTELP 180
```

```
Qy 181 AGLNGLENLDTLLQENSLYTIIPKGFSGSHLLPFAFLHGNPWLNCNCEILYFRRLQDNA 240
Db 181 AGLNGLENLDTLLQENSLYTIIPKGFSGSHLLPFAFLHGNPWLNCNCEILYFRRLQDNA 240
Qy 241 ENVYWKQVVDKAVTSNVASVQCDNSDKFPVYKYPKGCGTLDGEGDGLDYYPBEDT 300
Db 241 ENVYWKQGVDAAMTSNVASVQCDNSDKFPVYKYPKGCGTLDGEGDGLDYYPBEDT 300
Qy 301 EGDKVRPHTPCPCPAPEALGAPSVFLPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFN 360
Db 301 EGDKVRPHTPCPCPAPEALGAPSVFLPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFN 360
Qy 361 WYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPVPKEITI 420
Db 361 WYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPVPKEITI 420
Qy 421 SKAGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPP 480
Db 421 SKAGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPP 480
Qy 481 VLDSGDSFFLYSKLTVDKSRWQOQGNVFSCSVMEALHNHYTQKSLSLSPGK 531
Db 481 VLDSGDSFFLYSKLTVDKSRWQOQGNVFSCSVMEALHNHYTQKSLSLSPGK 531
```

RESULT 6

```
US-10-382-758-6
; Sequence 6, Application US/10382758
; Publication No. US20030232047A1
; GENERAL INFORMATION:
; APPLICANT: Shaw, Gray D.
; APPLICANT: Sako, Dianne S.
; APPLICANT: Kumar, Ravindra
; APPLICANT: Sullivan, Francis
; APPLICANT: McDonagh, Tom
; TITLE OF INVENTION: PlatiLet Glycoprotein IB Alpha Fusion Polypeptides and
; FILE REFERENCE: 22058-503
; CURRENT APPLICATION NUMBER: US/10/382,758
; PRIOR FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: US/10/068,426
; PRIOR FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/266,838
; PRIOR FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 531
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(531)
; OTHER INFORMATION: GP1b290/1A-Ig
US-10-382-758-6
```

```
Query Match 99.4%; Score 2823; DB 15; Length 531;
Best Local Similarity 99.4%; Pred. No. 1.2e-199;
Matches 528; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MPLLLLLLPSLPHPHPICEVSKVASHLEVNCCKRNLTALPPDLPKDTTILHSENLLY 60
Db 1 MPLLLLLLPSLPHPHPICEVSKVASHLEVNCCKRNLTALPPDLPKDTTILHSENLLY 60
Qy 61 TFSLATLMPYTRLTQNLNDRCELTKLQVDGTLPLVGLTDLSHNQLSLPLLGQTLPALTV 120
Db 61 TFSLATLMPYTRLTQNLNDRCELTKLQVDGTLPLVGLTDLSHNQLSLPLLGQTLPALTV 120
Qy 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSIANNLTLP 180
Db 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSIANNLTLP 180
```

```
Qy 181 AGLNGLENLDTLLQENSLYTIIPKGFSGSHLLPFAFLHGNPWLNCNCEILYFRRLQDNA 240
Db 181 AGLNGLENLDTLLQENSLYTIIPKGFSGSHLLPFAFLHGNPWLNCNCEILYFRRLQDNA 240
Qy 241 ENVYWKQVVDKAVTSNVASVQCDNSDKFPVYKYPKGCGTLDGEGDGLDYYPBEDT 300
Db 241 ENVYWKQGVDAAMTSNVASVQCDNSDKFPVYKYPKGCGTLDGEGDGLDYYPBEDT 300
Qy 301 EGDKVRPHTPCPCPAPEALGAPSVFLPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFN 360
Db 301 EGDKVRPHTPCPCPAPEALGAPSVFLPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFN 360
Qy 361 WYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPVPKEITI 420
Db 361 WYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPVPKEITI 420
Qy 421 SKAGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPP 480
Db 421 SKAGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPP 480
Qy 481 VLDSGDSFFLYSKLTVDKSRWQOQGNVFSCSVMEALHNHYTQKSLSLSPGK 531
Db 481 VLDSGDSFFLYSKLTVDKSRWQOQGNVFSCSVMEALHNHYTQKSLSLSPGK 531
```

RESULT 7

```
US-10-068-426-3
; Sequence 3, Application US/10068426
; Publication No. US20030091576A1
; GENERAL INFORMATION:
; APPLICANT: Shaw, Gray D.
; APPLICANT: Sako, Dianne S.
; APPLICANT: Kumar, Ravindra
; APPLICANT: Sullivan, Francis
; APPLICANT: McDonagh, Tom
; TITLE OF INVENTION: PlatiLet Glycoprotein IB Alpha Fusion Polypeptides and
; FILE REFERENCE: 22058-503
; CURRENT APPLICATION NUMBER: US/10/068,426
; PRIOR FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/266,838
; PRIOR FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(544)
; OTHER INFORMATION: GP1b302/4X-Ig
US-10-068-426-3
```

```
Query Match 99.4%; Score 2822.5; DB 14; Length 544;
Best Local Similarity 97.6%; Pred. No. 1.4e-199;
Matches 531; Conservative 0; Mismatches 0; Indels 13; Gaps 1;
```

```
Qy 1 MPLLLLLLPSLPHPHPICEVSKVASHLEVNCCKRNLTALPPDLPKDTTILHSENLLY 60
Db 1 MPLLLLLLPSLPHPHPICEVSKVASHLEVNCCKRNLTALPPDLPKDTTILHSENLLY 60
Qy 61 TFSLATLMPYTRLTQNLNDRCELTKLQVDGTLPLVGLTDLSHNQLSLPLLGQTLPALTV 120
Db 61 TFSLATLMPYTRLTQNLNDRCELTKLQVDGTLPLVGLTDLSHNQLSLPLLGQTLPALTV 120
Qy 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSIANNLTLP 180
Db 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSIANNLTLP 180
Qy 181 AGLNGLENLDTLLQENSLYTIIPKGFSGSHLLPFAFLHGNPWLNCNCEILYFRRLQDNA 240
Db 181 AGLNGLENLDTLLQENSLYTIIPKGFSGSHLLPFAFLHGNPWLNCNCEILYFRRLQDNA 240
```

```
QY 241 ENYVWKQVVDVAVTSNVAQVQVNDKFPVYKYPKGCPTLGDGDTLDLYDYYPEEDT 300
DB 241 ENYVWKQVVDVAVTSNVAQVQVNDKFPVYKYPKGCPTLGDGDTLDLYDYYPEEDT 300
QY 301 EGDQV-----RPHCTPCPAPEALGAPSVFLFPKPKDQTLMSRTPVETCVV 347
DB 301 EGDQVAAATATVVKFPTKARPHCTPCPAPEALGAPSVFLFPKPKDQTLMSRTPVETCVV 360
QY 348 DVSHEDPEVKFNWYVDGVEVHNATKPREEQNSTYRVVSVLTVLHQDWLNGKEYKCKV 407
DB 361 DVSHEDPEVKFNWYVDGVEVHNATKPREEQNSTYRVVSVLTVLHQDWLNGKEYKCKV 420
QY 408 SNKALPVPTEKTSKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWES 467
DB 421 SNKALPVPTEKTSKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWES 480
QY 468 NGQPENNYKTTTPVLDSDGFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNHYTQKSLSL 527
DB 481 NGQPENNYKTTTPVLDSDGFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNHYTQKSLSL 540
QY 528 SPQK 531
DB 541 SPQK 544
```

RESULT 8

```
US-10-382-758-3
; Sequence 3, Application US/10382758
; Publication No. US20030232047A1
; GENERAL INFORMATION:
; APPLICANT: Shaw, Gray D.
; APPLICANT: Sako, Dianne S.
; APPLICANT: Kumar, Ravindra
; APPLICANT: Sullivan, Francis
; APPLICANT: McDonagh, Tom
; TITLE OF INVENTION: Platellet Glycoprotein IB Alpha Fusion Polypeptides and
; TITLE OF INVENTION: Methods of Use Thereof
; FILE REFERENCE: 22058-503
; CURRENT APPLICATION NUMBER: US/10/382,758
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: US/10/068,426
; PRIOR FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/266,838
; PRIOR FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(544)
; OTHER INFORMATION: GP1b302/4X-Ig
US-10-382-758-3
```

Query Match 99.4%; Score 2822.5; DB 15; Length 544;
Best Local Similarity 97.6%; Pred. No. 1.4e-199;
Matches 531; Conservative 0; Mismatches 0; Indels 13; Gaps 1;

```
QY 1 MPLLLLLLLLPSLPHPICEVSKVASHLEVNCDEKNTALPPDLKDTTILHLSNLLY 60
DB 1 MPLLLLLLLLPSLPHPICEVSKVASHLEVNCDEKNTALPPDLKDTTILHLSNLLY 60
QY 61 TFSLATMPYTRLTQNLDRCELTQLQVDTGLPVLTGLDLSHNQLOSLPLGQTLPALTV 120
DB 61 TFSLATMPYTRLTQNLDRCELTQLQVDTGLPVLTGLDLSHNQLOSLPLGQTLPALTV 120
QY 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNNTLP 180
DB 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNNTLP 180
```

```
QY 181 AGLLNGLENLDTLLLOENSLYTIIPKGFPGSHLLPPAFLLHGNPWLNCCEILYFRRLWLDNA 240
DB 181 AGLLNGLENLDTLLLOENSLYTIIPKGFPGSHLLPPAFLLHGNPWLNCCEILYFRRLWLDNA 240
QY 241 ENYVWKQVVDVAVTSNVAQVQVNDKFPVYKYPKGCPTLGDGDTLDLYDYYPEEDT 300
DB 241 ENYVWKQVVDVAVTSNVAQVQVNDKFPVYKYPKGCPTLGDGDTLDLYDYYPEEDT 300
QY 301 EGDQV-----RPHCTPCPAPEALGAPSVFLFPKPKDQTLMSRTPVETCVV 347
DB 301 EGDQVAAATATVVKFPTKARPHCTPCPAPEALGAPSVFLFPKPKDQTLMSRTPVETCVV 360
QY 348 DVSHEDPEVKFNWYVDGVEVHNATKPREEQNSTYRVVSVLTVLHQDWLNGKEYKCKV 407
DB 361 DVSHEDPEVKFNWYVDGVEVHNATKPREEQNSTYRVVSVLTVLHQDWLNGKEYKCKV 420
QY 408 SNKALPVPTEKTSKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWES 467
DB 421 SNKALPVPTEKTSKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWES 480
QY 468 NGQPENNYKTTTPVLDSDGFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNHYTQKSLSL 527
DB 481 NGQPENNYKTTTPVLDSDGFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNHYTQKSLSL 540
QY 528 SPQK 531
DB 541 SPQK 544
```

RESULT 9

```
US-10-068-426-1
; Sequence 1, Application US/10068426
; Publication No. US20030091576A1
; GENERAL INFORMATION:
; APPLICANT: Shaw, Gray D.
; APPLICANT: Sako, Dianne S.
; APPLICANT: Kumar, Ravindra
; APPLICANT: Sullivan, Francis
; APPLICANT: McDonagh, Tom
; TITLE OF INVENTION: Platellet Glycoprotein IB Alpha Fusion Polypeptides and
; TITLE OF INVENTION: Methods of Use Thereof
; FILE REFERENCE: 22058-503
; CURRENT APPLICATION NUMBER: US/10/068,426
; CURRENT FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/266,838
; PRIOR FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(544)
; OTHER INFORMATION: GP1b302-Ig
US-10-068-426-1
```

Query Match 99.1%; Score 2812.5; DB 14; Length 544;
Best Local Similarity 97.2%; Pred. No. 7.4e-199;
Matches 529; Conservative 1; Mismatches 1; Indels 13; Gaps 1;

```
QY 1 MPLLLLLLLLPSLPHPICEVSKVASHLEVNCDEKNTALPPDLKDTTILHLSNLLY 60
DB 1 MPLLLLLLLLPSLPHPICEVSKVASHLEVNCDEKNTALPPDLKDTTILHLSNLLY 60
QY 61 TFSLATMPYTRLTQNLDRCELTQLQVDTGLPVLTGLDLSHNQLOSLPLGQTLPALTV 120
DB 61 TFSLATMPYTRLTQNLDRCELTQLQVDTGLPVLTGLDLSHNQLOSLPLGQTLPALTV 120
QY 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNNTLP 180
DB 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNNTLP 180
```

```
QY 181 AGLINGLENLDTLQLQENSLYTIIPKGFSGSHLLPFAFLHGNPWLNCNCEILYFRRLWQDNA 240
Db 181 AGLINGLENLDTLQLQENSLYTIIPKGFSGSHLLPFAFLHGNPWLNCNCEILYFRRLWQDNA 240
QY 241 ENVVYWKQVDVKAVTSNVASVQCDNSDKFPVYKYPKGCGCTLGDGDTLDLYDYYPEBDT 300
Db 241 ENVVYWKQVDVKAVTSNVASVQCDNSDKFPVYKYPKGCGCTLGDGDTLDLYDYYPEBDT 300
QY 301 EGDKV-----RPHTCPPCAPEALGAPSVLFPKPKDITLMISRTPEVTCV 347
Db 301 EGDKVATATVVKPTKARPHTCPPCAPEALGAPSVLFPKPKDITLMISRTPEVTCV 360
QY 348 VDVSHEDEPVKFNWYVDGVEVHNATKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKV 407
Db 361 VDVSHEDEPVKFNWYVDGVEVHNATKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKV 420
QY 408 SNKALPVPPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWES 467
Db 421 SNKALPVPPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWES 480
QY 468 NGQPENNYKTPPVLDSDGSEFFLYSKLTVDKSRWQOGNVFSCSVMHAEALHNHYTQKSLSL 527
Db 481 NGQPENNYKTPPVLDSDGSEFFLYSKLTVDKSRWQOGNVFSCSVMHAEALHNHYTQKSLSL 540
QY 528 SPKG 531
Db 541 SPKG 544

RESULT 10
US-10-068-426-2
; Sequence 2, Application US/10068426
; Publication No. US20030091576A1
; GENERAL INFORMATION:
; APPLICANT: Shaw, Gray D.
; APPLICANT: Sako, Dianne S.
; APPLICANT: Kumar, Ravindra
; APPLICANT: Sullivan, Francis
; APPLICANT: McDonagh, Tom
; TITLE OF INVENTION: Platiect Glycoprotein IB Alpha Fusion Polypeptides and
; TITLE OF INVENTION: Methods of Use Thereof
; FILE REFERENCE: 22058-503
; CURRENT APPLICATION NUMBER: US/10/068,426
; PRIOR FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/266,838
; PRIOR FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(544)
; OTHER INFORMATION: GPlb302/2A-Ig
US-10-068-426-2

Query Match 99.1%; Score 2812.5; DB 14; Length 544;
Best Local Similarity 97.2%; Pred. No. 7.4e-199;
Matches 529; Conservative 1; Mismatches 1; Indels 13; Gaps 1;

QY 1 MLLLLLLLLPSPLHPHPICEVSKVASHLEWNCNKRNLTAIPDLPKDITLHSENLLY 60
Db 1 MLLLLLLLLPSPLHPHPICEVSKVASHLEWNCNKRNLTAIPDLPKDITLHSENLLY 60
QY 61 TFSLATLMPYTRLTQNLDRCELTKLQVDGTLPLVGLTDLSHNQLSLPLLGQTLPALTV 120
Db 61 TFSLATLMPYTRLTQNLDRCELTKLQVDGTLPLVGLTDLSHNQLSLPLLGQTLPALTV 120
QY 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGCLLTPPKLEKLSLANNNTLPL 180
```

```
Db 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGCLLTPPKLEKLSLANNNTLPL 180
QY 181 AGLINGLENLDTLQLQENSLYTIIPKGFSGSHLLPFAFLHGNPWLNCNCEILYFRRLWQDNA 240
Db 181 AGLINGLENLDTLQLQENSLYTIIPKGFSGSHLLPFAFLHGNPWLNCNCEILYFRRLWQDNA 240
QY 241 ENVVYWKQVDVKAVTSNVASVQCDNSDKFPVYKYPKGCGCTLGDGDTLDLYDYYPEBDT 300
Db 241 ENVVYWKQVDVKAVTSNVASVQCDNSDKFPVYKYPKGCGCTLGDGDTLDLYDYYPEBDT 300
QY 301 EGDKV-----RPHTCPPCAPEALGAPSVLFPKPKDITLMISRTPEVTCV 347
Db 301 EGDKAATATVVKPTKARPHTCPPCAPEALGAPSVLFPKPKDITLMISRTPEVTCV 360
QY 348 VDVSHEDEPVKFNWYVDGVEVHNATKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKV 407
Db 361 VDVSHEDEPVKFNWYVDGVEVHNATKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKV 420
QY 408 SNKALPVPPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWES 467
Db 421 SNKALPVPPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWES 480
QY 468 NGQPENNYKTPPVLDSDGSEFFLYSKLTVDKSRWQOGNVFSCSVMHAEALHNHYTQKSLSL 527
Db 481 NGQPENNYKTPPVLDSDGSEFFLYSKLTVDKSRWQOGNVFSCSVMHAEALHNHYTQKSLSL 540
QY 528 SPKG 531
Db 541 SPKG 544

RESULT 11
US-10-382-758-1
; Sequence 1, Application US/10382758
; Publication No. US20030232047A1
; GENERAL INFORMATION:
; APPLICANT: Shaw, Gray D.
; APPLICANT: Sako, Dianne S.
; APPLICANT: Kumar, Ravindra
; APPLICANT: Sullivan, Francis
; APPLICANT: McDonagh, Tom
; TITLE OF INVENTION: Platiect Glycoprotein IB Alpha Fusion Polypeptides and
; TITLE OF INVENTION: Methods of Use Thereof
; FILE REFERENCE: 22058-503
; CURRENT APPLICATION NUMBER: US/10/382,758
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: US/10/068,426
; PRIOR FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/266,838
; PRIOR FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(544)
; OTHER INFORMATION: GPlb302-Ig
US-10-382-758-1

Query Match 99.1%; Score 2812.5; DB 15; Length 544;
Best Local Similarity 97.2%; Pred. No. 7.4e-199;
Matches 529; Conservative 1; Mismatches 1; Indels 13; Gaps 1;

QY 1 MLLLLLLLLPSPLHPHPICEVSKVASHLEWNCNKRNLTAIPDLPKDITLHSENLLY 60
Db 1 MLLLLLLLLPSPLHPHPICEVSKVASHLEWNCNKRNLTAIPDLPKDITLHSENLLY 60
QY 61 TFSLATLMPYTRLTQNLDRCELTKLQVDGTLPLVGLTDLSHNQLSLPLLGQTLPALTV 120
Db 61 TFSLATLMPYTRLTQNLDRCELTKLQVDGTLPLVGLTDLSHNQLSLPLLGQTLPALTV 120
```

```
QY 121 LDVSFNRRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPPKLEKSLANNLTLP 180
| | | | |
DB 121 LDVSFNRRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPPKLEKSLANNLTLP 180
| | | | |
QY 181 AGLLNGLENLDTLLLOENSLYTIIPKGFSGSHLLPFAFLHGNPWLNCCEILYFRWLQDNA 240
| | | | |
DB 181 AGLLNGLENLDTLLLOENSLYTIIPKGFSGSHLLPFAFLHGNPWLNCCEILYFRWLQDNA 240
| | | | |
QY 241 ENVYVMKQVVDVAVTSNVASVQCDNSDKFPVYKPKGCGPTLGDGDTLDLYYPEEDT 300
| | | | |
DB 241 ENVYVMKQVVDVAVTSNVASVQCDNSDKFPVYKPKGCGPTLGDGDTLDLYYPEEDT 300
| | | | |
QY 301 EGDKV-----RPHCTPPCPAPEALGAPSVFLPPPKDKTLMISRTPEVTCV 347
| | | | |
DB 301 EGDKVRAIRTVKFPKARPHCTPPCPAPEALGAPSVFLPPPKDKTLMISRTPEVTCV 360
| | | | |
QY 348 VDVSHEDEPEVKFNWYVDGVEVHNKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKV 407
| | | | |
DB 361 VDVSHEDEPEVKFNWYVDGVEVHNKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKV 420
| | | | |
QY 408 SNKALPVPTEKTSKAKGQPREPVYTLPPSREMTKNQVSLTCLVKGFYPSDIAVEWES 467
| | | | |
DB 421 SNKALPVPTEKTSKAKGQPREPVYTLPPSREMTKNQVSLTCLVKGFYPSDIAVEWES 480
| | | | |
QY 468 NGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQOGNPFSCSMHEALHNHYTQKSLSL 527
| | | | |
DB 481 NGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQOGNPFSCSMHEALHNHYTQKSLSL 540
| | | | |
QY 528 SPQK 531
| | | | |
DB 541 SPQK 544
```

RESULT 12

```
US-10-382-758-2
; Sequence 2, Application US/10382758
; Publication No. US20030232047A1
; GENERAL INFORMATION:
; APPLICANT: Shaw, Gray D.
; APPLICANT: Sako, Dianne S.
; APPLICANT: Kumar, Ravindra
; APPLICANT: Sullivan, Francis
; APPLICANT: McDonagh, Tom
; TITLE OF INVENTION: Platellet Glycoprotein IB Alpha Fusion Polypeptides and
; TITLE OF INVENTION: Methods of Use Thereof
; FILE REFERENCE: 22058-503
; CURRENT APPLICATION NUMBER: US/10/382,758
; PRIOR FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: US/10/068,426
; PRIOR FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/266,838
; PRIOR FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(544)
; OTHER INFORMATION: GP1b302/2A-Ig
US-10-382-758-2
```

```
Query Match 99.1%; Score 2812.5; DB 15; Length 544;
Best Local Similarity 97.2%; Pred. No. 7.4e-199;
Matches 529; Conservative 1; Mismatches 1; Indels 13; Gaps 1;
QY 1 MPELLLLLLPSLPHPICEVSKVASHLEVNCCKENLTALPPDLPKDTTILHLSNLLY 60
| | | | |
DB 1 MPELLLLLLPSLPHPICEVSKVASHLEVNCCKENLTALPPDLPKDTTILHLSNLLY 60
| | | | |
```

```
QY 61 TFSLATLMPYTRLTQNLNDRCELTKLQVDGTLPVLGTLDSLHNQLOSLPLLGOTLPALT 120
| | | | |
DB 61 TFSLATLMPYTRLTQNLNDRCELTKLQVDGTLPVLGTLDSLHNQLOSLPLLGOTLPALT 120
| | | | |
QY 121 LDVSFNRRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPPKLEKSLANNLTLP 180
| | | | |
DB 121 LDVSFNRRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPPKLEKSLANNLTLP 180
| | | | |
QY 181 AGLLNGLENLDTLLLOENSLYTIIPKGFSGSHLLPFAFLHGNPWLNCCEILYFRWLQDNA 240
| | | | |
DB 181 AGLLNGLENLDTLLLOENSLYTIIPKGFSGSHLLPFAFLHGNPWLNCCEILYFRWLQDNA 240
| | | | |
QY 241 ENVYVMKQVVDVAVTSNVASVQCDNSDKFPVYKPKGCGPTLGDGDTLDLYYPEEDT 300
| | | | |
DB 241 ENVYVMKQVVDVAVTSNVASVQCDNSDKFPVYKPKGCGPTLGDGDTLDLYYPEEDT 300
| | | | |
QY 301 EGDKV-----RPHCTPPCPAPEALGAPSVFLPPPKDKTLMISRTPEVTCV 347
| | | | |
DB 301 EGDKVAATATVVKFPKARPHCTPPCPAPEALGAPSVFLPPPKDKTLMISRTPEVTCV 360
| | | | |
QY 348 VDVSHEDEPEVKFNWYVDGVEVHNKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKV 407
| | | | |
DB 361 VDVSHEDEPEVKFNWYVDGVEVHNKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKV 420
| | | | |
QY 408 SNKALPVPTEKTSKAKGQPREPVYTLPPSREMTKNQVSLTCLVKGFYPSDIAVEWES 467
| | | | |
DB 421 SNKALPVPTEKTSKAKGQPREPVYTLPPSREMTKNQVSLTCLVKGFYPSDIAVEWES 480
| | | | |
QY 468 NGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQOGNPFSCSMHEALHNHYTQKSLSL 527
| | | | |
DB 481 NGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQOGNPFSCSMHEALHNHYTQKSLSL 540
| | | | |
QY 528 SPQK 531
| | | | |
DB 541 SPQK 544
```

RESULT 13

```
US-10-408-765A-430
; Sequence 430, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 430
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-430
```

```
Query Match 56.9%; Score 1615.5; DB 16; Length 626;
Best Local Similarity 84.2%; Pred. No. 1.7e-110;
Matches 319; Conservative 5; Mismatches 20; Indels 35; Gaps 5;
```

```
QY 1 MPELLLLLLPSLPHPICEVSKVASHLEVNCCKENLTALPPDLPKDTTILHLSNLLY 60
| | | | |
DB 1 MPELLLLLLPSLPHPICEVSKVASHLEVNCCKENLTALPPDLPKDTTILHLSNLLY 60
| | | | |
QY 61 TFSLATLMPYTRLTQNLNDRCELTKLQVDGTLPVLGTLDSLHNQLOSLPLLGOTLPALT 120
| | | | |
DB 61 TFSLATLMPYTRLTQNLNDRCELTKLQVDGTLPVLGTLDSLHNQLOSLPLLGOTLPALT 120
| | | | |
```

```
QY 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTLP 180
Db 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTLP 180
QY 181 AGLINGLENLDTLLLOENSLYTI PKGFFGSHLLPPAFHGNPWLNCNCEILYFRWLQDNA 240
Db 181 AGLINGLENLDTLLLOENSLYTI PKGFFGSHLLPPAFHGNPWLNCNCEILYFRWLQDNA 240
QY 241 ENVVYWKQVDVKAVTSNVASVQCDNSDKFPVYKPGKCGPTLGDGDTLDYDYPEEDT 300
Db 241 ENVVYWKQVDVKAVTSNVASVQCDNSDKFPVYKPGKCGPTLGDGDTLDYDYPEEDT 300
QY 301 EGDKVR-----PHTCP-----PCPA---PEALGAPSVFLFPPK- - 330
Db 301 EGDKVRATRVVVKPTKAHTTPWGLFYSWSTASLDSQMPSSLHPTQESTKEQTTFPPRWT 360
QY 331 PKDTL-----MISRTPEVT 344
Db 361 PNFTLHMSITFSKTPKST 379

RESULT 14
US-10-741-600-1496
; Sequence 1496, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; CURRENT APPLICATION NUMBER: CLO01499
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1496
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-600-1496

Query Match 56.9%; Score 1615.5; DB 17; Length 626;
Best Local Similarity 84.2%; Pred. No. 1.7e-110;
Matches 319; Conservative 5; Mismatches 20; Indels 35; Gaps 5;

QY 1 MPLLALLLPSPLPHPHICVSKVASHLEVNCCKRNLTALPPDLPKDPTTLHLSENLLY 60
Db 1 MPLLALLLPSPLPHPHICVSKVASHLEVNCCKRNLTALPPDLPKDPTTLHLSENLLY 60
QY 61 TFSLATLMPYTRLTQNLDRCELTKLQVDGTLPLVGLTDLDSHNQLSLPLLGQTLPALTV 120
Db 61 TFSLATLMPYTRLTQNLDRCELTKLQVDGTLPLVGLTDLDSHNQLSLPLLGQTLPALTV 120
QY 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTLP 180
Db 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTLP 180
QY 181 AGLINGLENLDTLLLOENSLYTI PKGFFGSHLLPPAFHGNPWLNCNCEILYFRWLQDNA 240
Db 181 AGLINGLENLDTLLLOENSLYTI PKGFFGSHLLPPAFHGNPWLNCNCEILYFRWLQDNA 240
QY 241 ENVVYWKQVDVKAVTSNVASVQCDNSDKFPVYKPGKCGPTLGDGDTLDYDYPEEDT 300
Db 241 ENVVYWKQVDVKAVTSNVASVQCDNSDKFPVYKPGKCGPTLGDGDTLDYDYPEEDT 300
QY 301 EGDKVR-----PHTCP-----PCPA---PEALGAPSVFLFPPK- - 330
Db 301 EGDKVRATRVVVKPTKAHTTPWGLFYSWSTASLDSQMPSSLHPTQESTKEQTTFPPRWT 360
QY 331 PKDTL-----MISRTPEVT 344
Db 361 PNFTLHMSITFSKTPKST 379
```

```
RESULT 15
US-10-068-426-11
; Sequence 11, Application US/10068426
; Publication No. US20030091576A1
; GENERAL INFORMATION:
; APPLICANT: Shaw, Gray D.
; APPLICANT: Sako, Dianne S.
; APPLICANT: Kumar, Ravindra
; APPLICANT: Sullivan, Francis
; APPLICANT: McDonagh, Tom
; TITLE OF INVENTION: Platiect Glycoprotein IB Alpha Fusion Polypeptides and
; FILE OF INVENTION: Methods of Use Thereof
; TITLE REFERENCE: 22058-503
; CURRENT APPLICATION NUMBER: US/10/068,426
; CURRENT FILING DATE: 2002-02-06
; PRIOR FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 290
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(290)
; OTHER INFORMATION: GBLb290/2V
US-10-068-426-11

Query Match 54.1%; Score 1536; DB 14; Length 290;
Best Local Similarity 100.0%; Pred. No. 4.5e-105;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 HPICEVSKVASHLEVNCCKRNLTALPPDLPKDPTTLHLSENLLYTFSLATLMPYTRLTOL 76
Db 1 HPICEVSKVASHLEVNCCKRNLTALPPDLPKDPTTLHLSENLLYTFSLATLMPYTRLTOL 60
QY 77 NLDRCCELTKLQVDGTLPLVGLTDLDSHNQLSLPLLGQTLPALTVLDVSNRLTSLPLGAL 136
Db 61 NLDRCCELTKLQVDGTLPLVGLTDLDSHNQLSLPLLGQTLPALTVLDVSNRLTSLPLGAL 120
QY 137 RGLGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTLPAGLNGLENLDTLLQ 196
Db 121 RGLGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTLPAGLNGLENLDTLLQ 180
QY 197 ENSLYTTPKGFPGSHLLPPAFHGNPWLNCNCEILYFRWLQDNAENVYWKQVDVKAVT 256
Db 181 ENSLYTTPKGFPGSHLLPPAFHGNPWLNCNCEILYFRWLQDNAENVYWKQVDVKAVT 240
QY 257 SNVASVQCDNSDKFPVYKPGKCGPTLGDGDTLDYDYPEEDTEGDKVR 306
Db 241 SNVASVQCDNSDKFPVYKPGKCGPTLGDGDTLDYDYPEEDTEGDKVR 290

Search completed: May 24, 2005, 06:16:23
Job time : 138 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2005, 05:54:11 ; Search time 42 Seconds
(without alignments)
943.777 Million cell updates/sec

Title: US-10-068-426-5

Perfect score: 2839

Sequence: 1 MFLLLLLLPSPHPHPC.....MHEALHNHYTKSLSPGK 531

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

1: /cgn2_6/prodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/prodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/prodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/prodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/prodata/1/iaa/PCFUS_COMB.pep.*
6: /cgn2_6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1528.5	53.8	320	1	US-07-613-083B-1
2	1521.5	53.6	610	1	US-07-821-717B-6
3	1521.5	53.6	610	1	US-08-119-262B-6
4	1521.5	53.6	610	1	US-08-135-929A-11
5	1521.5	53.6	610	1	US-08-234-265A-11
6	1237	43.6	313	3	US-08-713-555F-36
7	1230	43.3	388	3	US-09-131-247-16
8	1230	43.3	388	4	US-09-784-623-16
9	1230	43.3	449	3	US-08-897-236-23
10	1230	43.3	449	4	US-09-500-253B-23
11	1226.5	43.2	397	4	US-08-775-066-2
12	1223	43.1	451	2	US-08-887-352B-14
13	1223	43.1	451	2	US-08-887-352B-16
14	1223	43.1	451	2	US-08-887-352B-18
15	1223	43.1	451	3	US-08-466-151-65
16	1223	43.1	451	3	US-09-109-207C-14
17	1223	43.1	451	3	US-09-109-207C-16
18	1223	43.1	451	3	US-09-109-207C-18
19	1223	43.1	451	3	US-09-282-505-2
20	1223	43.1	451	3	US-09-054-255-2
21	1223	43.1	451	3	US-09-296-005-14
22	1223	43.1	451	3	US-09-296-005-16
23	1223	43.1	451	3	US-09-296-005-18
24	1223	43.1	451	4	US-09-282-846-2
25	1223	43.1	451	4	US-09-680-145-2
26	1223	43.1	451	4	US-09-920-171-14
27	1223	43.1	451	4	US-09-920-171-16

28 1223 43.1 451 4 US-09-920-171-18 Sequence 18, Appl
29 1223 43.1 451 4 US-09-716-028-14 Sequence 14, Appl
30 1223 43.1 451 4 US-09-716-028-16 Sequence 16, Appl
31 1223 43.1 451 4 US-09-716-028-18 Sequence 18, Appl
32 1223 43.1 451 4 US-09-483-588-2 Sequence 2, Appl
33 1223 43.1 451 4 US-10-113-996-14 Sequence 14, Appl
34 1223 43.1 451 4 US-10-113-996-16 Sequence 16, Appl
35 1223 43.1 451 4 US-10-113-996-18 Sequence 18, Appl
36 1222 43.0 453 3 US-08-466-151-8 Sequence 8, Appl
37 1222 43.0 453 3 US-08-466-163B-8 Sequence 8, Appl
38 1222 43.0 453 4 US-09-802-096-8 Sequence 8, Appl
39 1222 43.0 453 4 US-09-802-077-8 Sequence 8, Appl
40 1222 43.0 977 4 US-09-590-656-1 Sequence 1, Appl
41 1222 43.0 977 4 US-09-733-764-1 Sequence 1, Appl
42 1220.5 43.0 694 4 US-09-313-942-18 Sequence 18, Appl
43 1220.5 43.0 793 4 US-09-313-942-32 Sequence 32, Appl
44 1220 43.0 631 4 US-09-056-461-22 Sequence 22, Appl
45 1218 42.9 452 4 US-09-773-877B-16 Sequence 16, Appl

ALIGNMENTS

RESULT 1
US-07-613-083B-1
; Sequence 1, Application US/07613083B
; Patent No. 5340727
; GENERAL INFORMATION:
; APPLICANT: Ruggeri, Zaverio M.
; APPLICANT: Ware, Jerry, inventors
; APPLICANT: on behalf of Scripps Clinic and Research
; APPLICANT: Foundation
; TITLE OF INVENTION: GPib' Fragments and Recombinant
; TITLE OF INVENTION: DNA Expression Vectors
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scripps Clinic and Research
; ADDRESSEE: Foundation
; STREET: 10666 No. 5340727th Torrey Pines Road
; CITY: La Jolla
; STATE: California
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette-5.25 inch, 1.2 Mb
; COMPUTER: AST Bravo IBM PC comp. (386SX)
; OPERATING SYSTEM: MS DOS version 3.2
; SOFTWARE: WordPerfect 5.1 conv. to ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/613,083B
; FILING DATE: 19911114
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA: This appl. is a c-i-p of
; APPLICATION NUMBER: U.S. 07/470,674
; FILING DATE: 04-Jan-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Barron, Alexis
; REGISTRATION NUMBER: 22,702
; REFERENCE/DOCKET NUMBER: P16,569-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 923-4466
; TELEFAX: (215) 923-2189
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 320
; TYPE: AMINO ACID
; STRANDEDNESS: No. 5340727 applicable
; TOPOLOGY: Linear
; US-07-613-083B-1

Query Match 53.8%; Score 1528.5; DB 1; Length 320;
Best Local Similarity 95.1%; Pred. No. 6.8e-129;
Matches 291; Conservative 1; Mismatches 3; Indels 11; Gaps 1;

QY 17 HPICEVSKVASHLEVNCDCRNLTPALPPDLPKDPTTLHLSENLLYTFSLATLMPYTRLTQ 76
Db 1 HPICEVSKVASHLEVNCDCRNLTPALPPDLPKDPTTLHLSENLLYTFSLATLMPYTRLTQ 60
QY 77 NLDRCCLTKLQVDTGLPVLGTLDSLHNOQLSLPLGQTLPALTVLDVSNFRLTSLPLGAL 136
Db 61 NLDRCCLTKLQVDTGLPVLGTLDSLHNOQLSLPLGQTLPALTVLDVSNFRLTSLPLGAL 120
QY 137 RGLGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLPAGLGLNGLENLDTLLQ 196
Db 121 RGLGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLPAGLGLNGLENLDTLLQ 180
QY 197 ENSLYTTPKGFPGSHLLPFAFLHGNPWLNCCEILYFRRLQDNAENVVVKQVVDVKAVT 256
Db 181 ENSLYTTPKGFPGSHLLPFAFLHGNPWLNCCEILYFRRLQDNAENVVVKQVVDVKAVT 240
QY 257 SNVASVQCDNSDKFPVVKYKPGKCPPTLGDEGDTLDLYYPEEDTEGDKVR 306
Db 241 SNVASVQCDNSDKFPVVKYKPGKCPPTLGDEGDTLDLYYPEEDTEGDKVRATRTVVKPPT 300
QY 307 -PHTCP 311
Db 301 KAHTTP 306

RESULT 2
US-07-821-717B-6
; Sequence 6, Application US/07821717B
; Patent No. 5298239
; GENERAL INFORMATION:
; APPLICANT: Miller, Jonathan L.
; APPLICANT: Cunningham, David
; APPLICANT: Lyle, Vicki A.
; APPLICANT: Finch, Clara N.
; TITLE OF INVENTION: MUTATIONS RENDERING PLATELET
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
; COUNTRY: USA
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/821,717B
; FILING DATE: 15-JAN-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Timain, Susan J.
; REGISTRATION NUMBER: 34,103
; REFERENCE/DOCKET NUMBER: 20884/21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 263-1636
; TELEFAX: (716) 263-1600
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 610 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PUBLICATION INFORMATION:
; AUTHORS: Lopez, Jose A.
; AUTHORS: Chung, Dominic W.
; AUTHORS: Fujikawa, Kazuo
; AUTHORS: Hagen, Frederick S.

; AUTHORS: Papayannopoulou, Thalia
; AUTHORS: Roth, Gerald J.
; TITLE: Cloning of the alpha chain of human
; TITLE: platelet glycoprotein Ib: A transmembrane protein with homology
; TITLE: to leucine-rich alpha-2-glycoprotein
; JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
; VOLUME: 84
; PAGES: 5615-5619
; DATE: AUG-1987
; RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 0 TO 610
; PUBLICATION INFORMATION:
; AUTHORS: Zimmerman, Theodore S.
; AUTHORS: Ruggeri, Zaverio M.
; AUTHORS: Houghten, Richard A.
; AUTHORS: Vincete, Vincete
; AUTHORS: Mohri, Hiroshi
; TITLE: Proteolytic fragments and synthetic
; TITLE: peptides that block the binding of von Willebrand factor to the
; TITLE: platelet membrane glycoprotein Ib
; DOCUMENT NUMBER: EP 0 317 278 A2
; FILING DATE: 16-NOV-1988
; PUBLICATION DATE: 24-MAY-1989
; RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 0 TO 293
; US-07-821-717B-6

Query Match 53.6%; Score 1521.5; DB 1; Length 610;
Best Local Similarity 82.9%; Pred. No. 7.8e-128;
Matches 301; Conservative 5; Mismatches 22; Indels 35; Gaps 5;
QY 17 HPICEVSKVASHLEVNCDCRNLTPALPPDLPKDPTTLHLSENLLYTFSLATLMPYTRLTQ 76
Db 1 HPICEVSKVASHLEVNCDCRNLTPALPPDLPKDPTTLHLSENLLYTFSLATLMPYTRLTQ 60
QY 77 NLDRCCLTKLQVDTGLPVLGTLDSLHNOQLSLPLGQTLPALTVLDVSNFRLTSLPLGAL 136
Db 61 NLDRCCLTKLQVDTGLPVLGTLDSLHNOQLSLPLGQTLPALTVLDVSNFRLTSLPLGAL 120
QY 137 RGLGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLPAGLGLNGLENLDTLLQ 196
Db 121 RGLGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLPAGLGLNGLENLDTLLQ 180
QY 197 ENSLYTTPKGFPGSHLLPFAFLHGNPWLNCCEILYFRRLQDNAENVVVKQVVDVKAVT 256
Db 181 ENSLYTTPKGFPGSHLLPFAFLHGNPWLNCCEILYFRRLQDNAENVVVKQVVDVKAVT 240
QY 257 SNVASVQCDNSDKFPVVKYKPGKCPPTLGDEGDTLDLYYPEEDTEGDKVR 306
Db 241 SNVASVQCDNSDKFPVVKYKPGKCPPTLGDEGDTLDLYYPEEDTEGDKVRATRTVVKPPT 300
QY 307 -PHTCP 311
Db 301 KAHTTPWGLFYSWSTASLDSQMPSSLHPTQESTKEQTTFPRWTPNFTLHMSITFSKTP 360

RESULT 3
US-08-119-262B-6
; Sequence 6, Application US/08119262B:
; Patent No. 5492809
; GENERAL INFORMATION:
; APPLICANT: Miller, Jonathan L.
; APPLICANT: Cunningham, David
; APPLICANT: Lyle, Vicki A.
; APPLICANT: Finch, Clara N.
; TITLE OF INVENTION: MUTATIONS RENDERING PLATELET
; TITLE OF INVENTION: GLYCOPROTEIN Ib ALPHA LESS REACTIVE
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle
; STREET: Clinton Square, P.O. Box 1051

CITY: Rochester
STATE: New York
COUNTRY: USA
ZIP: 14603
COMPUTER: IBM PC compatible
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/119,262B
FILING DATE: 09-SEP-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/821,717
FILING DATE: 15-JAN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Timain, Susan J.
REGISTRATION NUMBER: 34,103
REFERENCE/DOCKET NUMBER: 20884/22
TELEPHONE: (716) 263-1636
TELEFAX: (716) 263-1600
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 610 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
PUBLICATION INFORMATION:
AUTHORS: Lopez, Jose A.
AUTHORS: Chung, Dominic W.
AUTHORS: Fujikawa, Kazuo
AUTHORS: Hagen, Frederick S.
AUTHORS: Papayannopoulou, Thalia
AUTHORS: Roth, Gerald J.
TITLE: Cloning of the alpha chain of human platelet glycoprotein Ib: A transmembrane protein
TITLE: leucine-rich alpha-2-glycoprotein
JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
VOLUME: 84
PAGES: 5615-5619
DATE: AUG-1987
RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 0 TO 610
PUBLICATION INFORMATION:
AUTHORS: Zimmerman, Theodore S.
AUTHORS: Ruggeri, Zaverio M.
AUTHORS: Houghten, Richard A.
AUTHORS: Vincete, Vincete
AUTHORS: Mohri, Hiroshi
TITLE: Proteolytic fragments and synthetic peptides
TITLE: that block the binding of von Willebrand factor to the membrane glycoprotein Ib
DOCUMENT NUMBER: EP 0 317 278 A2
FILING DATE: 16-NOV-1988
PUBLICATION DATE: 24-MAY-1989
RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 0 TO 293
US-08-119-262B-6

Query Match 53.6%; Score 1521.5; DB 1; Length 610;
Best Local Similarity 82.9%; Pred. No. 7.8e-128;
Matches 301; Conservative 5; Mismatches 22; Indels 35; Gaps 5;
QY 17 HPICEVSKVASHLEVNCNKLTPDLPKDTIHLSENLYTFSLATLMPYTRLTQ 76
DB 1 HPICEVSKVASHLEVNCNKLTPDLPKDTIHLSENLYTFSLATLMPYTRLTQ 60
QY 77 NLDRCETLKQVDCGTLPLVGLTDLSHNQLSLPLGOTLPALTVLDVSNRLTSLPLGAL 136
DB 61 NLDRCETLKQVDCGTLPLVGLTDLSHNQLSLPLGOTLPALTVLDVSNRLTSLPLGAL 120
QY 137 RGLGELQELYKGNELKTLPPGLLTPPKLEKLSLANNLTLPAGLLNGLENLDTLLQ 196

DB 121 RGLGELQELYKGNELKTLPPGLLTPPKLEKLSLANNLTLPAGLLNGLENLDTLLQ 180
QY 197 ENSLYTIPKGFSGSHLLPFAFLHGNPWLNCNCEILYFRRWLQDQNAENVYVWKQVVDVKAVT 256
DB 181 ENSLYTIPKGFSGSHLLPFAFLHGNPWLNCNCEILYFRRWLQDQNAENVYVWKQVVDVKAVT 240
QY 257 SNVASVQCDNSDKFPVYKYPGKCPITLGDGDTLDLYYPEEDTEGDKVR----- 306
DB 241 SNVASVQCDNSDKFPVYKYPGKCPITLGDGDTLDLYYPEEDTEGDKVRATRVVKFPT 300
QY 307 -PHTCP-----PCPA---PEALGAPSVFLPPK--PKDTL-----MISRT 341
DB 301 KAHTPWGLFYSWSTASLDSQWESSLHPTQESTKEQTTTPPRWTPNFTLHESITFSKTP 360
QY 342 EVT 344
DB 361 KST 363
RESULT 4
US-08-135-929A-11
Sequence 11, Application US/08135929A
Patent No. 5593959
GENERAL INFORMATION:
APPLICANT: Miller, Jonathan L.
APPLICANT: Cunningham, David
APPLICANT: Lyle, Vicki A.
APPLICANT: Finch, Clara N.
APPLICANT: Pincus, Matthew R.
TITLE OF INVENTION: Mutations in the Gene Encoding the Alpha
TITLE OF INVENTION: Chain of Platelet Glycoprotein Ib
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon, Hargrave, Devans & Doyle
STREET: Clinton Square, P.O. Box 1051
CITY: Rochester
STATE: New York
COUNTRY: USA
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/135,929A
FILING DATE: 14-OCT-1993
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Timian, Susan J.
REGISTRATION NUMBER: 34,103
REFERENCE/DOCKET NUMBER: 20884/23
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 263-1636
TELEFAX: (716) 263-1600
TELEX: 978450
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 610 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-135-929A-11

Query Match 53.6%; Score 1521.5; DB 1; Length 610;
Best Local Similarity 82.9%; Pred. No. 7.8e-128;
Matches 301; Conservative 5; Mismatches 22; Indels 35; Gaps 5;
QY 17 HPICEVSKVASHLEVNCNKLTPDLPKDTIHLSENLYTFSLATLMPYTRLTQ 76
DB 1 HPICEVSKVASHLEVNCNKLTPDLPKDTIHLSENLYTFSLATLMPYTRLTQ 60

QY 77 NLDRCETKQLQVDTLPVLGTLDSLHNOQLSLPLGQTLPALTVLDVSFNRLTSLPLGAL 136
DB 61 NLDRCETKQLQVDTLPVLGTLDSLHNOQLSLPLGQTLPALTVLDVSFNRLTSLPLGAL 120
QY 137 RGLGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTTELPAGLLNGLENLDTLLQ 196
DB 121 RGLGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTTELPAGLLNGLENLDTLLQ 180
QY 197 ENSLYTTPKGFPGSHLLPFAFLHGNPWLNCCEILYFRRLWQDNNAENVVWQGVVDVKAVT 256
DB 181 ENSLYTTPKGFPGSHLLPFAFLHGNPWLNCCEILYFRRLWQDNNAENVVWQGVVDVKAVT 240
QY 257 SNVASVQCDNSDKFPVYKPGKCGPTLGDGDTLDLYYPEEDTEGDKVR----- 306
DB 241 SNVASVQCDNSDKFPVYKPGKCGPTLGDGDTLDLYYPEEDTEGDKVRATRTVVKFPT 300
QY 307 -PHTCP-----PCPA---PEALGAPSVFLFPK--PKDTL-----MISRTP 341
DB 301 KAHTTPWGLFYSWSTASLDSQMPSSLHPTQBSTKEQTFPPRPWTFNFTLHMESITFSKTP 360
QY 342 EVT 344
DB 361 KST 363

RESULT 5

US-08-234-265A-11
; Sequence 11, Application US/08234265A
; Patent No. 5624817
; GENERAL INFORMATION:
; APPLICANT: Miller, Jonathan L.
; APPLICANT: Cunningham, David
; APPLICANT: Lyle, Vicki A.
; APPLICANT: Finch, Clara N.
; APPLICANT: Pincus, Matthew R.
; TITLE OF INVENTION: Mutations in the Gene Encoding the
; TITLE OF INVENTION: Chain of Platelet Glycoprotein Ib
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
; COUNTRY: USA
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/234,265A
; FILING DATE: 28-APR-1994
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Timian, Susan J.
; REGISTRATION NUMBER: 34,103
; REFERENCE/DOCKET NUMBER: 20884/24
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 263-1636
; TELEFAX: (716) 263-1600
; TELEX: 978450
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 610 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-234-265A-11

Query Match 53.6%; Score 1521.5; DB 1; Length 610;
Best Local Similarity 82.9%; Pred. No. 7.8e-128;
Matches 301; Conservative 5; Mismatches 22; Indels 35; Gaps 5;

QY 17 HPICEVSKVASHLEVNCDCRNLTALEPDLPKDTTILHLENLTYTFSLATLMPYTRLTOL 76
DB 1 HPICEVSKVASHLEVNCDCRNLTALEPDLPKDTTILHLENLTYTFSLATLMPYTRLTOL 60
QY 77 NLDRCETKQLQVDTLPVLGTLDSLHNOQLSLPLGQTLPALTVLDVSFNRLTSLPLGAL 136
DB 61 NLDRCETKQLQVDTLPVLGTLDSLHNOQLSLPLGQTLPALTVLDVSFNRLTSLPLGAL 120
QY 137 RGLGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTTELPAGLLNGLENLDTLLQ 196
DB 121 RGLGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTTELPAGLLNGLENLDTLLQ 180
QY 197 ENSLYTTPKGFPGSHLLPFAFLHGNPWLNCCEILYFRRLWQDNNAENVVWQGVVDVKAVT 256
DB 181 ENSLYTTPKGFPGSHLLPFAFLHGNPWLNCCEILYFRRLWQDNNAENVVWQGVVDVKAVT 240
QY 257 SNVASVQCDNSDKFPVYKPGKCGPTLGDGDTLDLYYPEEDTEGDKVR----- 306
DB 241 SNVASVQCDNSDKFPVYKPGKCGPTLGDGDTLDLYYPEEDTEGDKVRATRTVVKFPT 300
QY 307 -PHTCP-----PCPA---PEALGAPSVFLFPK--PKDTL-----MISRTP 341
DB 301 KAHTTPWGLFYSWSTASLDSQMPSSLHPTQBSTKEQTFPPRPWTFNFTLHMESITFSKTP 360
QY 342 EVT 344
DB 361 KST 363

RESULT 6

US-08-713-556F-36
; Sequence 36, Application US/08713556F
; Patent No. 6277975
; GENERAL INFORMATION:
; APPLICANT: Larsen, Glenn
; APPLICANT: Sako, Dianne
; APPLICANT: Chang, Xiao Jia
; APPLICANT: Veldman, Geertuida M.
; APPLICANT: Cumming, Dale
; APPLICANT: Kumar, Ravindra
; APPLICANT: Shaw, Gray
; TITLE OF INVENTION: NOVEL P-SELECTIN LIGAND PROTEIN
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LEGAL AFFAIRS
; STREET: 87 CAMBRIDGE PARK DRIVE
; CITY: CAMBRIDGE
; STATE: MA
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/713,556F
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/965,662
; FILING DATE: 23-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/112,608
; FILING DATE: 26-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/10168
; FILING DATE: 22-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/235,398
; FILING DATE: 28-APR-1994
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/316,305
FILING DATE: 30-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/428,734
FILING DATE: 25-APR-1995
ATTORNEY/AGENT INFORMATION:
NAME: BROWN, SCOTT A.
REGISTRATION NUMBER: 32,724
REFERENCE/DOCKET NUMBER: GI 5213F
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8224
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 313 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-713-556F-36

Query Match 43.6%; Score 1237; DB 3; Length 313;
Best Local Similarity 97.0%; Pred. No. 1e-102;
Matches 229; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 296 PESTTEGDKVRPHTCPPCPAPEALGAPSVLPFPKPKDTLMISRTPEVTCVVVDVSHEDP 355
DB 78 PESTTEVPAARPHTCPPCPAPEALGAPSVLPFPKPKDTLMISRTPEVTCVVVDVSHEDP 137

QY 356 EVKFNWYDGVVHNAKTPREBOYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPVP 415
DB 138 EVKFNWYDGVVHNAKTPREBOYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPVP 197

QY 416 IEKTIKAKGQPREPOVYTLPPSRREMTKNQVSLTCLVKGFPYSDIAVEWESNGQPENNY 475
DB 198 IEKTIKAKGQPREPOVYTLPPSRREMTKNQVSLTCLVKGFPYSDIAVEWESNGQPENNY 257

QY 476 KITPPVLDSDGSPFLYKSLTVDKSRWQOGNPFVSCVWMEALHNHYTKLSLSLSPGK 531
DB 258 KITPPVLDSDGSPFLYKSLTVDKSRWQOGNPFVSCVWMEALHNHYTKLSLSLSPGK 313

RESULT 7
US-09-131-247-16
; Sequence 16, Application US/09131247
; Patent No. 6294170
; GENERAL INFORMATION:
; APPLICANT: Boone, Thomas C.
; APPLICANT: Hershenson, Susan
; APPLICANT: Bevilacqua, Michael P.
; APPLICANT: Collins, David S.
; TITLE OF INVENTION: COMPOSITION AND METHOD FOR TREATING INFLAMMATORY
; TITLE OF INVENTION: DISEASES
; FILE REFERENCE: A-365F
; CURRENT APPLICATION NUMBER: US/09/131,247
; CURRENT FILING DATE: 1998-08-07
; EARLIER APPLICATION NUMBER: 60/055,185
; EARLIER FILING DATE: 1997-08-08
; EARLIER APPLICATION NUMBER: PCT/US 97/02131
; EARLIER FILING DATE: 1997-02-10
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 388
; TYPE: PRT
; ORGANISM: Human
US-09-131-247-16

Query Match 43.3%; Score 1230; DB 3; Length 388;
Best Local Similarity 64.5%; Pred. No. 5.9e-102;
Matches 249; Conservative 20; Mismatches 69; Indels 48; Gaps 6;

QY 171 LANNLTLPAGLNG-----LENLDTLILQENSLYTIKPGFPGSHLLPFAFLHGNFWLC 225
DB 26 LRNN---QLVAGYLOGPNVNLKIDVVPFIEPHAL-----FLGIHGGMCL 68

QY 226 NCEILYFRRLQDQNAENVYVWKQVVDVKAFTSNVASVQCDNSDKFPVYKPGKCP--- 281
DB 69 SCVKSGETRLQLEAVN-----ITDLSNRKQDKRFAFIRSDSGTTTSFESAACFGWFL 122

QY 282 -----TLGDEGDTLDYIYPEDTEGD---KVRPHTCPPCPAPEALGAPSVF 325
DB 123 CTAMEADQPVSLTNMPDEGVMTKPYFQDEAAABPKSDKTKHTCCPCPAPELLGGPSVF 182

QY 326 LPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYDGVVHNAKTPREBOYNSTYR 385
DB 183 LPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYDGVVHNAKTPREBOYNSTYR 242

QY 386 VVSVLTVLHQDWLNGKEYCKVSNKALPVPVPIETKISKAKGQPREPOVYTLPPSRREMTKN 445
DB 243 VVSVLTVLHQDWLNGKEYCKVSNKALPVPVPIETKISKAKGQPREPOVYTLPPSRDELTKN 302

QY 446 QVSLTCLVKGFPYSDIAVEWESNGQPENNYKTTTPVLDSDGSPFLYKSLTVDKSRWQOGN 505
DB 303 QVSLTCLVKGFPYSDIAVEWESNGQPENNYKTTTPVLDSDGSPFLYKSLTVDKSRWQOGN 362

QY 506 VFSCSVMEALHNHYTKLSLSLSPGK 531
DB 363 VFSCSVMEALHNHYTKLSLSLSPGK 388

Db 26 LRNN---QLVAGYLOGPNVNLKIDVVPFIEPHAL-----FLGIHGGMCL 68
QY 226 NCEILYFRRLQDQNAENVYVWKQVVDVKAFTSNVASVQCDNSDKFPVYKPGKCP--- 281
Db 69 SCVKSGETRLQLEAVN-----ITDLSNRKQDKRFAFIRSDSGTTTSFESAACFGWFL 122
QY 282 -----TLGDEGDTLDYIYPEDTEGD---KVRPHTCPPCPAPEALGAPSVF 325
Db 123 CTAMEADQPVSLTNMPDEGVMTKPYFQDEAAABPKSDKTKHTCCPCPAPELLGGPSVF 182
QY 326 LPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYDGVVHNAKTPREBOYNSTYR 385
Db 183 LPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYDGVVHNAKTPREBOYNSTYR 242
QY 386 VVSVLTVLHQDWLNGKEYCKVSNKALPVPVPIETKISKAKGQPREPOVYTLPPSRREMTKN 445
Db 243 VVSVLTVLHQDWLNGKEYCKVSNKALPVPVPIETKISKAKGQPREPOVYTLPPSRDELTKN 302
QY 446 QVSLTCLVKGFPYSDIAVEWESNGQPENNYKTTTPVLDSDGSPFLYKSLTVDKSRWQOGN 505
Db 303 QVSLTCLVKGFPYSDIAVEWESNGQPENNYKTTTPVLDSDGSPFLYKSLTVDKSRWQOGN 362
QY 506 VFSCSVMEALHNHYTKLSLSLSPGK 531
Db 363 VFSCSVMEALHNHYTKLSLSLSPGK 388

RESULT 8
US-09-784-623-16
; Sequence 16, Application US/09784623
; Patent No. 6733753
; GENERAL INFORMATION:
; APPLICANT: Boone, Thomas C.
; APPLICANT: Hershenson, Susan
; APPLICANT: Bevilacqua, Michael P.
; APPLICANT: Collins, David S.
; TITLE OF INVENTION: COMPOSITION AND METHOD FOR TREATING INFLAMMATORY
; TITLE OF INVENTION: DISEASES
; FILE REFERENCE: A-365F
; CURRENT APPLICATION NUMBER: US/09/784,623
; CURRENT FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: 09/131,247
; PRIOR FILING DATE: 1998-08-07
; PRIOR APPLICATION NUMBER: PCT/US 97/02131
; PRIOR FILING DATE: 1997-02-10
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 388
; TYPE: PRT
; ORGANISM: Human
US-09-784-623-16

Query Match 43.3%; Score 1230; DB 4; Length 388;
Best Local Similarity 64.5%; Pred. No. 5.9e-102;
Matches 249; Conservative 20; Mismatches 69; Indels 48; Gaps 6;

QY 171 LANNLTLPAGLNG-----LENLDTLILQENSLYTIKPGFPGSHLLPFAFLHGNFWLC 225
Db 26 LRNN---QLVAGYLOGPNVNLKIDVVPFIEPHAL-----FLGIHGGMCL 68

QY 226 NCEILYFRRLQDQNAENVYVWKQVVDVKAFTSNVASVQCDNSDKFPVYKPGKCP--- 281
Db 69 SCVKSGETRLQLEAVN-----ITDLSNRKQDKRFAFIRSDSGTTTSFESAACFGWFL 122

QY 282 -----TLGDEGDTLDYIYPEDTEGD---KVRPHTCPPCPAPEALGAPSVF 325
Db 123 CTAMEADQPVSLTNMPDEGVMTKPYFQDEAAABPKSDKTKHTCCPCPAPELLGGPSVF 182

QY 326 LPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYDGVVHNAKTPREBOYNSTYR 385
Db 183 LPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYDGVVHNAKTPREBOYNSTYR 242

RESULT 10	US-09-500-253B-23	Sequence 23, Application US/09500253B	Patent No. 6500640	GENERAL INFORMATION:	APPLICANT: Regeneron Pharmaceuticals, Inc.	TITLE OF INVENTION: Modified Oral Tissue Affecting Factor and Composition	FILE REFERENCE: REG 133-Z	CURRENT APPLICATION NUMBER: US/09/500,253B	CURRENT FILING DATE: 2000-02-08	NUMBER OF SEQ ID NOS: 27	SOFTWARE: PatentIn version 3.0	SEQ ID NO 23	LENGTH: 449	TYPE: PRT	ORGANISM: Human	US-09-500-253B-23
	Query Match	43.3%;	Score 1230;	DB 4;	Length 449;											
	Best Local Similarity	56.1%;	Pred. No. 7.4e-102;													
	Matches 273;	Conservative 21;	Mismatches 95;	Indels 98;	Gaps 13;											
QY	105	LQSLPLLGQTLPALTVLDVFNRLTSLPLGLRGL														139
DB	1	MERCPSLGVITYALVV														56
QY	140	--CELOELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTELPAAGLNGLENLDTLLLOE														197
DB	57	KEKDLNETLRLSLGGHYDFGWATSPEDRPGGGG														111
QY	198	NSLYTIPKPGFGSHLLPFAFLHG														249
DB	112	-----PSGAMPSEIKLEFSEGLAQGLQWLWSQTECPVLY--AW--NDLGSRFWPRY														160
QY	250	VDV														289
DB	161	VKVGCSCFKRSCVPEGMVCKPSKSHLTVLRWCQ														209
QY	290	DLVDYTP----	BEDTEGDKVRPHTCCPPCAPEALGAPSVFLPPPKDITLMISRTPEVT													344
DB	210	-----YPIISECKSCSGDKT--HTCDPCPAPELLGGPSVFLPPPKDITLMISRTPEVT														262
QY	345	CVVVDVSHEDPEVKFNWYDVGVEVHNAKTPREEQYNSTYRVVSVLTVLHODWLNKGEYK														404
DB	263	CVVVDVSHEDPEVKFNWYDVGVEVHNAKTPREEQYNSTYRVVSVLTVLHODWLNKGEYK														322
QY	405	CKVSNKALVPVIEKTTISKAKQPREPOVYTLPPSREEMTKNQVSLTCLVKGFFYPSDIAVE														464
DB	323	CKVSNKALPAPIEKTTISKAKQPREPOVYTLPPSREEMTKNQVSLTCLVKGFFYPSDIAVE														382
QY	465	WESNGOPENNYKTTPPVLDSGDSGFYLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTOKS														524
DB	383	WESNGOPENNYKTTPPVLDSGDSGFYLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTOKS														442
QY	525	LSLSPGK 531														
DB	443	LSLSPGK 449														
RESULT 11	US-08-775-066-2	Sequence 2, Application US/08775066	Patent No. 6620413	GENERAL INFORMATION:	APPLICANT: Desauvage, Frederic	TITLE OF INVENTION: Ob protein-immunoglobulin chimeras	NUMBER OF SEQUENCES: 2	CORRESPONDENCE ADDRESS:	ADDRESSEE: Genentech, Inc.	STREET: 460 Point San Bruno Blvd	CITY: South San Francisco	STATE: California				

```

; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/775,066
; FILING DATE: 27-Dec-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: P0985R1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-3216
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 397 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; US-08-775-066-2

Query Match 43.2%; Score 1226.5; DB 4; Length 397;
Best Local Similarity 66.0%; Pred. No. 1.3e-101;
Matches 256; Conservative 17; Mismatches 48; Indels 67; Gaps 9;

QY 157 PGL---LTPPKLEKLSLANNITELPA-----GLNGLENLDTLLLOENSLYTIKGFPG 209
DB 64 PGLHPTLTLSKMDQTLAVYQQLTSPSRNVIQISNDLENLRDL----- 108

QY 210 SHLLPFAFLHGNPWLNCETLYFRWLQDAENVY---VWKQVDVKAVTSNVASV--Q 263
DB 109 -HVIAFSKCHLPWASGLETL-----DSLGVLEASGYSTVEVVALSLQSLQMLWQ 160

QY 264 CONSDKFPVYKYPKGKPTLGDGDTLDYDYPEEDTEGDKVRPHTCPCPAPEALGAPS 323
DB 161 LDLS-----PGC-----GVTD-----KTHCTPCPAPELLGGPS 189

QY 324 VFLFPKPKDITLMISTPTEVTCVVVDVSHEDPVRKNNWYVDGVEVHNATKPRREQYNST 383
DB 190 VFLFPKPKDITLMISTPTEVTCVVVDVSHEDPVRKNNWYVDGVEVHNATKPRREQYNST 249

QY 384 YRVVSVLTVLHODWLNKGYKCKVSNKALPVPPIEKTISKAKGQPREPOVYITLPPSREMT 443
DB 250 YRVVSVLTVLHODWLNKGYKCKVSNKALPVPPIEKTISKAKGQPREPOVYITLPPSREMT 309

QY 444 KQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSGFELYSKLTVDKSRWQ 503
DB 310 KQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSGFELYSKLTVDKSRWQ 369

QY 504 GNVFSCVMHEALHNHYTQKSLSLSPGK 531
DB 370 GNVFSCVMHEALHNHYTQKSLSLSPGK 397

RESULT 12
US-08-887-352B-14
; Sequence 14, Application US/08887352B
; Patent No. 5994511
; GENERAL INFORMATION:
; APPLICANT: Henry B. Lowman, Leonard G. Presta, Paula M. Jardieu, John Lowe
; TITLE OF INVENTION: Improved Anti-IgE Antibodies and Method of
; TITLE OF INVENTION: Improving Polypeptides
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA

```

```

; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/887,352B
; FILING DATE: 03-Jul-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Svoboda, Craig G.
; REGISTRATION NUMBER: 39,044
; REFERENCE/DOCKET NUMBER: P1123
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1489
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 451 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; US-08-887-352B-14

Query Match 43.1%; Score 1223; DB 2; Length 451;
Best Local Similarity 64.1%; Pred. No. 3.2e-101;
Matches 254; Conservative 14; Mismatches 58; Indels 70; Gaps 10;

QY 188 ENLDLTLLOENSLYTIKGF-----GSHLLPFAFLHGNPWLNCETLYFRWLQDAENV 243
DB 74 DSKNTFYLOMNSLRABDTAVYICARGSHY-----FGH---W-----HFAVMGQGLTVT 119

QY 244 -----YVWKQVDVKAVTSNVASVQCDNSDKPP-----VVKYPGKGC 280
DB 120 SSASTKGPSVFLPAPSSKSTSGTALGCLVDYFPPEPTVSVNSGALITGVHTFPA--- 176

QY 281 PTLGDEG-----DTDLY-----DYYPEEDTEGDKVRP-----HTCPCPA 315
DB 177 -VLQSSGLSYLSVTVTPSSSLGTQTYICNVNHPKSNKTVDKKVPKSCDKTHTCPCPA 235

QY 316 PEALGAPSVLFPKPKDITLMISTPTEVTCVVVDVSHEDPVRKNNWYVDGVEVHNATK 375
DB 236 PELLGGPSVFLFPKPKDITLMISTPTEVTCVVVDVSHEDPVRKNNWYVDGVEVHNATK 295

QY 376 REEQYNSTYRVVSVLTVLHODWLNKGYKCKVSNKALPVPPIEKTISKAKGQPREPOVYITL 435
DB 296 REEQYNSTYRVVSVLTVLHODWLNKGYKCKVSNKALPVPPIEKTISKAKGQPREPOVYITL 355

QY 436 PPSREMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSGFELYSKLT 495
DB 356 PPSREMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSGFELYSKLT 415

QY 496 VDKSRWQQGNVFCSCVMHEALHNHYTQKSLSLSPGK 531
DB 416 VDKSRWQQGNVFCSCVMHEALHNHYTQKSLSLSPGK 451

RESULT 13
US-08-887-352B-16
; Sequence 16, Application US/08887352B
; Patent No. 5994511
; GENERAL INFORMATION:
; APPLICANT: Henry B. Lowman, Leonard G. Presta, Paula M. Jardieu, John Lowe
; TITLE OF INVENTION: Improved Anti-IgE Antibodies and Method of
; TITLE OF INVENTION: Improving Polypeptides
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA

```



```

; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,151
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/466163
; FILING DATE: 06-Jun-1995
; APPLICATION NUMBER: 08/405617
; FILING DATE: 15-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/185899
; FILING DATE: 26-JAN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/879495
; FILING DATE: 07-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/744768
; FILING DATE: 14-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Svoboda, Craig G.
; REGISTRATION NUMBER: 39,044
; REFERENCE/DOCKET NUMBER: P0718P2C1D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1489
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 451 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-08-466-151-65

```

Query Match	43.1%;	Score 1223;	DB 3;	Length 451;		
Best Local Similarity	64.1%;	Pred. No. 3.2e-101;				
Matches 254;	Conservative 14;	Mismatches 58;	Indels 70;	Gaps 10;		
QY	188	ENLDTLLQNSLYTI	PKGFF----	QSHLLPPAPFLHGNPWL	CNCEILYFRRLQD	NAENV 243
DB	74	DSKNTFYLOWNSL	RAEDTAVY	CARGSHY----	FGH----	W-----HFAVWGCGTLVTV 119
QY	244	-----YVMQVVDV	KAVTSNV	VSQCNSDKFP	-----	VYKYPKGC 280
DB	120	SSASTKGPSVF	PLAPSSKTS	GGTAALGCLVKDY	FPPEPVT	VSWSNGALTSGVHTFPA--- 176
QY	281	PTLGDG----	-----	DTDLV----	DYYPEEDTEGDKVP	-----HTCPCPA 315
DB	177	VLQSSGLYSLSS	VTVVSSSLGTQT	YICNNHNPSTK	VDKVKPEKSCDK	TKTHCPCPA 235
QY	316	PEALGAPSVFL	PPPKPKD	TLMISRTPEVTC	VVDVSHEDPEV	FNFWTVDGVGVHNAKTCP 375
DB	236	PELLGGSPVFL	PPPKPKD	TLMISRTPEVTC	VVDVSHEDPEV	FNFWTVDGVGVHNAKTCP 295
QY	376	RSQYNSTYRW	SVSLTVLHOD	WLNGKEYKCK	SVSNKALP	VPDIKTTISAKAQPREPQVYTL 435
DB	296	RSQYNSTYRW	SVSLTVLHOD	WLNGKEYKCK	SVSNKALP	VPDIKTTISAKAQPREPQVYTL 355
QY	436	PPSREEMTKNQ	VSILTC	LVKGFYPSDIA	VENESNGOPEN	NYKTTTPPVLDSDGSPFLYSKLT 495
DB	356	PPSREEMTKNQ	VSILTC	LVKGFYPSDIA	VENESNGOPEN	NYKTTTPPVLDSDGSPFLYSKLT 415
QY	496	VDKSRWQGNV	FGSCVM	HEALHNNHY	TKO	SLSPGK 531
DB	416	VDKSRWQGNV	FGSCVM	HEALHNNHY	TKO	SLSPGK 451

Search completed: May 24, 2005, 06:09:49
Job time : 44 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2005, 05:54:10 ; Search time 42 Seconds
(without alignments)
1216.455 Million cell updates/sec

Title: US-10-068-426-5
Perfect score: 2839
Sequence: 1 MFLLLLLLLPSLPHHPIC.....MHEALHNHYTKSLSPGK 531
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1615.5	56.9	626	1 NBHUIA	platelet glycoprot
2	1205	42.4	330	1 GHU	Ig gamma-1 chain C
3	1203	42.4	374	2 S69339	Ig heavy chain V x
4	1197	42.2	255	4 S31866	Ig gamma-1 chain C x
5	1152	40.6	234	2 PT0207	Ig gamma chain C x
6	1148	40.4	326	1 G2HU	Ig gamma-2 chain C
7	1145	40.3	377	2 A23511	Ig gamma-3 chain C
8	1143	40.3	377	2 A60764	Ig gamma-3 chain C
9	1134.5	40.0	327	1 G4HU	Ig gamma-4 chain C
10	1116.5	39.3	289	1 G3H0WI	Ig gamma-3 heavy c
11	921	32.4	333	1 GHRB	Ig gamma chain C x
12	913	32.2	328	2 I47160	Ig gamma 2b chain
13	913	32.2	328	2 I47159	Ig gamma 2a chain
14	906	31.9	277	2 I47162	Ig gamma 4 chain c
15	893.5	31.5	328	2 I47158	Ig gamma 1 chain c
16	884.5	31.2	328	1 G2GP	Ig gamma 1 chain C
17	881	31.0	328	2 I47161	Ig gamma 3 chain C
18	857.5	30.2	470	2 S22080	Ig gamma 3 chain c
19	849.5	29.9	308	2 C30554	Ig heavy chain pre
20	849.5	29.9	472	2 S31459	Ig heavy chain C x
21	841	29.6	329	1 G3MSC	Ig gamma-1 chain -
22	838.5	29.5	444	2 PC4436	Ig gamma-3 chain C
23	834.5	29.4	326	2 PS0017	Ig gamma-1 chain C
24	830	29.2	398	1 G3MSM	Ig gamma-3 chain C
25	826.5	29.1	324	1 G1MS	Ig gamma-1 chain C
26	826.5	29.1	333	2 PS0018	Ig gamma-2b chain
27	821.5	28.9	329	2 S00847	Ig gamma-2c chain
28	821.5	28.9	393	1 G1MSM	Ig gamma-1 chain C
29	817.5	28.8	322	2 PS0019	Ig gamma-2a chain

RESULT 1

NBHUIA

platelet glycoprotein Ib alpha chain precursor - human
N;Alternate names: membrane glycoprotein Ib alpha chain

C;Contains: glycosialicin

C;Species: Homo sapiens (man)

C;Date: 28-Dec-1987 #sequence revision 28-Dec-1987 #text change 09-Jul-2004

C;Accession: A94174; A60435; A94173; S16945; I55355; A27075; A27102

R;Lopez, J.A.; Chung, D.W.; Fujikawa, K.; Hagen, F.S.; Papayannopoulou, T.; Roth, G.J.

Proc. Natl. Acad. Sci. U.S.A. 84, 5615-5619, 1987

A;Title: Cloning of the alpha-chain of human platelet glycoprotein Ib: a transmembrane p

A;Reference number: A94174; MUID:87289655; PMID:3303030

A;Accession: A94174

A;Molecule type: mRNA

A;Residues: 1-626 <LOP>

A;Cross-references: UNIPROT:P07359; GB:J02940; NID:G183499; PIDN:AAAS5595.1; PID:G306793

R;Wicki, A.N.; Walz, A.; Gerber-Huber, S.N.; Wenger, R.H.; Vornhagen, R.; Clemetson, K.J

Thromb. Haemost. 61, 448-453, 1989

A;Title: Isolation and characterization of human blood platelet mRNA and construction of
d cloning of a GPIb coding cDNA insert.

A;Reference number: A60435; MUID:90020160; PMID:2799758

A;Accession: A60435

A;Molecule type: mRNA

A;Residues: 207-467 <WIC>

R;Titani, K.; Takio, K.; Handa, M.; Ruggeri, Z.M.

Proc. Natl. Acad. Sci. U.S.A. 84, 5610-5614, 1987

A;Title: Amino acid sequence of the von Willebrand factor-binding domain of platelet mem

A;Reference number: A94173; MUID:87289654; PMID:3497398

A;Accession: A94173

A;Molecule type: protein

A;Residues: 17-315 <TIT>

R;Hess, D.; Schaller, J.; Rickli, E.E.; Clemetson, K.J.

Eur. J. Biochem. 199, 389-393, 1991

A;Title: Identification of the disulphide bonds in human platelet glycosialicin.

A;Reference number: S16945; MUID:91301149; PMID:2070794

A;Accession: S16945

A;Status: preliminary

A;Molecule type: protein

A;Residues: 224-227;262-270;277-282 <HES>

R;Lopez, J.A.; Ludwig, E.H.; McCarthy, B.J.

J. Biol. Chem. 267, 10055-10061, 1992

A;Title: Polymorphism of human glycoprotein Ib alpha results from a variable number of t
ations.

A;Reference number: I55355; MUID:92250564; PMID:1577776

A;Accession: I55355

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 412-427 <RES>

A;Cross-references: GB:S34436; NID:G249176; PIDN:AB22152.1; PID:G249177

A;Note: variant D

C;Comment: Glycoprotein Ib (GPIb), a surface membrane protein of platelets, participates

C;Comment: Platelet activation apparently involves disruption of the macromolecular comp

C;Comment: Binding sites for von Willebrand factor and thrombin (the latter site with un
C;Comment: Glycocalicin, which is approximately coextensive with the extracellular part
C;Genetics:
A;Gene: GDB:GPIBA; GPIB
A;Cross-references: GDB:118806; OMIM:231200
A;Map position: 1pter-17p12
C;Complex: heterodimer with platelet glycoprotein Ib beta chain (NBHU1B)
C;Superfamily: platelet glycoprotein Ib alpha chain; leucine-rich alpha-2-glycoprotein I
C;Keywords: blood coagulation; duplication; glycoprotein; platelet membrane; tandem repe
F:1-16/Domain: signal sequence #status predicted <SIG>
F:17-626/Product: platelet glycoprotein Ib alpha chain #status predicted <MPT>
F:148-71/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F:72-93/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F:94-116/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F:117-140/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
F:141-164/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
F:165-188/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
F:189-212/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>
F:379-430/Region: proline/threonine-rich 9-residue repeats
F:502-540/Domain: transmembrane #status predicted <TRM>
F:541-626/Domain: intracellular #status predicted <INT>
F:37.175/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:308/Binding site: carbohydrate (Thr) (covalent) #status experimental

Query Match 56.9%; Score 1615.5; DB 1; Length 626;
Best Local Similarity 84.2%; Pred. No. 5.3e-101;
Matches 319; Conservative 5; Mismatches 20; Indels 35; Gaps 5;

Qy	1	MPLLLLLLPSPLPHPCISVSKVASHLEVNCNRLTALPPDLPKDTTILHSENLLY 60
Db	1	MPLLLLLLPSPLPHPCISVSKVASHLEVNCNRLTALPPDLPKDTTILHSENLLY 60
Qy	61	TFSLATLMPYRLTQNLDRCLTKLQVDGTLPLVGLTDLGSHNLOSLPLGLQTLPALTV 120
Db	61	TFSLATLMPYRLTQNLDRCLTKLQVDGTLPLVGLTDLGSHNLOSLPLGLQTLPALTV 120
Qy	121	LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSANNLTPL 180
Db	121	LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSANNLTPL 180
Qy	181	AGLLNGLENLTLLLOENSLYTIKGFPGSHLLPFAFLHGNPWLNCBILYFRWLQDNA 240
Db	181	AGLLNGLENLTLLLOENSLYTIKGFPGSHLLPFAFLHGNPWLNCBILYFRWLQDNA 240
Qy	241	ENVVYWKQVDVKAVTSNVASVQCDNSDKFPVYKPGKCTPLGDEGTDLDYYPEDT 300
Db	241	ENVVYWKQVDVKAVTSNVASVQCDNSDKFPVYKPGKCTPLGDEGTDLDYYPEDT 300
Qy	301	EGDKVR-----PHTCP-----PCPA-----PEALGAPSVFLFPKP-- 330
Db	301	EGDKVRATRTVVKPTKAHTTPWGLFYGSWSTASLDSPSSLIHPQESTKEQTTFPPRWT 360
Qy	331	PKDTL-----MISRTPEVT 344
Db	361	PNFTLHMESITTSKTPKST 379

RESULT 2
GHU
Ig gamma-1 chain C region - human
C;Species: Homo sapiens (man)
C;Date: 31-Jan-1981 #sequence revision 18-Aug-1982 #text_change 09-Jul-2004
C;Accession: A93433; S36861; S33887; B90563; A90564; B91668; A91723; A02146
R;Ellison, J.W.; Berson, B.J.; Hood, L.E.
Nucleic Acids Res. 10, 4071-4079, 1982
A;Title: The nucleotide sequence of a human immunoglobulin C-gamma1 gene.
A;Reference number: A93433; MUID:82274238; PMID:6287432
A;Accession: A93433
A;Molecule type: DNA
A;Residues: 1-330 <ELL>
A;Cross-references: UNIPROT:P01857; EMBL:Z17370
A;Note: This sequence has the Gln(17) allotypic marker, 97-Lys, and the Gln(1) markers,
A;Note: Lys-330 is removed after translation

R;Harris, L.J.
submitted to the EMBL Data Library, October 1992
A;Reference number: S33904
A;Accession: S36861
A;Molecule type: DNA
A;Residues: 2-330 <HAR>
A;Cross-references: EMBL:Z17370
R;Takahashi, N.; Ueda, S.; Obata, M.; Nikaido, T.; Nakai, S.; Honjo, T.
Cell 29, 671-679, 1982
A;Title: Structure of human immunoglobulin gamma genes: implications for evolution of a
A;Reference number: S33887; MUID:83001943; PMID:6811139
A;Accession: S33887
A;Molecule type: DNA
A;Residues: 88-113;235-330 <TAK>
R;Cunningham, B.A.; Rutishauser, U.; Gall, W.E.; Gottlieb, P.D.; Waxdal, M.J.; Edelman,
Biochemistry 9, 3161-3170, 1970
A;Title: The covalent structure of a human gammaG-immunoglobulin. VII. Amino acid sequen
A;Reference number: A90563; MUID:71064024; PMID:5489771
A;Contents: myeloma protein Eu
A;Accession: B90563
A;Molecule type: protein
A;Residues: 1-96, R', 98-135 <CUN>
R;Rutishauser, U.; Cunningham, B.A.; Bennett, C.; Konigsberg, W.H.; Edelman, G.M.
Biochemistry 9, 3171-3181, 1970
A;Title: The covalent structure of a human gammaG-immunoglobulin. VIII. Amino acid sequen
A;Reference number: A90564; MUID:71064025; PMID:5530842
A;Contents: Eu
A;Accession: A90564
A;Molecule type: protein
A;Residues: 136-154, 'Q', 156-165, 'Q', 167-176, 'Q', 178-194, 'N', 196-197, 'D', 199-238, 'E', 240,
R;Ponstingl, H.; Hilschmann, N.
Hoppe-Seyler's Z. Physiol. Chem. 357, 1571-1604, 1976
A;Title: Die Primaerstruktur eines monoklonalen IgG1-Immunglobulins (Myelomprotein Nite),
igen Primaerstruktur.
A;Reference number: A91668; MUID:77070269; PMID:826475
A;Contents: myeloma protein Nie
A;Accession: B91668
A;Molecule type: protein
A;Residues: 1-34, 'Q', 36-96, 'K', 98-115, 'Q', 117-197, 'D', 199-238, 'D', 240, 'L', 242-268, 'E', 27
R;Schmidt, W.E.; Jung, H.D.; Palm, W.; Hilschmann, N.
Hoppe-Seyler's Z. Physiol. Chem. 364, 713-747, 1983
A;Title: Die Primaerstruktur des kristallisierten monoklonalen Immunglobulins IgG1 KOL
A;Reference number: A91723; MUID:83289131; PMID:6884994
A;Contents: myeloma protein KOL; disulfide bonds
A;Accession: A91723
A;Molecule type: protein
A;Residues: 1-96, 'R', 98-197, 'D', 199-238, 'E', 240, 'M', 242-266, 'D', 268-271, 'D', 273-330 <SCH
R;Gall, W.E.; Edelman, G.M.
Biochemistry 9, 3188-3196, 1970
A;Title: The covalent structure of a human gammaG-immunoglobulin. X. Intrachain disulfid
A;Reference number: A90565; MUID:71064027; PMID:4923144
A;Contents: annotation; disulfide bonds
R;Draker, L.; Schwarz, J.; Reichel, W.; Hilschmann, N.
Hoppe-Seyler's Z. Physiol. Chem. 357, 1515-1540, 1976
A;Title: Rule of antibody structure. The primary structure of monoclonal IgG1 immunoglob
enbromide cleavage products, and the disulfide bridges.
A;Reference number: A91667; MUID:77070267; PMID:1002129
A;Contents: annotation; disulfide bonds
C;Genetics:
A;Gene: GDB:IGHG1
A;Cross-references: GDB:120085; OMIM:147100
A;Map position: 14q32.33-14q32.33
A;Introns: 99/1; 114/1; 224/1
C;Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la
C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F:20-85/Domain: immunoglobulin homology <IMI>

F:137-206/Domain: immunoglobulin homology <IM2>
F:243-310/Domain: immunoglobulin homology <IM3>
F:27-83,144-204,250-308/Distulfide bonds: #status experimental
F:103/Distulfide bonds: interchain (to light chain) #status experimental
F:109,112/Distulfide bonds: interchain (to heavy chain) #status experimental
F:180/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 42.4%; Score 1205; DB 1; Length 330;
Best Local Similarity 73.7%; Pred. No. 9.5e-74;
Matches 235; Conservative 10; Mismatches 30; Indels 44; Gaps 5;

QY 253 KAVTSNVASVQCDNSDKP-----VVKPKGCGPTLGDEG----- 287
DB 16 KSTSGTALGCLVKDYFPEPTVTVSNVSGALTSQVHTFPA----VLQSGLSVLSVWTV 71
QY 288 -----DITDLY----DYYPEDTEGDKVRP-----HTCCPPCPAPEALGAPSVFLFPKPK 332
DB 72 PSSSLGTQYICNVNPKPNTKVDKVKPKCDKTHCCPPAPEALGAPSVFLFPKPK 131
QY 333 DTLMSRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTKPREQYNSTYRVVSVLTV 392
DB 132 DTLMSRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTKPREQYNSTYRVVSVLTV 191
QY 393 LHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRREMTKNQVSLTCL 452
DB 192 LHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCL 251
QY 453 VKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFPLYSKLTVDKSRWQQGNVFCVM 512
DB 252 VKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFPLYSKLTVDKSRWQQGNVFCVM 311
QY 513 HEALHNHYTQKSLSLSPGK 531
DB 312 HEALHNHYTQKSLSLSPGK 330

RESULT 3
Ig heavy chain V region precursor - human
C:Species: Homo sapiens (man)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 01-Dec-2000
C:Accession: S69339; S72664
R:Khamlichi, A.A.; Aucouturier, P.; Preud'homme, J.L.; Cogne, M.
Eur. J. Biochem. 229, 54-60, 1995
A:Title: Structure of abnormal heavy chains in human heavy-chain-deposition disease.
A:Reference number: S69339; MUID:95262687; PMID:7744049
A:Accession: S69339
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-374 <KHA>
A:Cross-references: EMBL:X81695
R:Khamlichi, A.A.
submitted to the EMBL Data Library, September 1994
A:Reference number: S72664
A:Accession: S72664
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-140, 'C', 142-374 <KH2>
A:Cross-references: EMBL:X81695
C:Superfamily: immunoglobulin C region; immunoglobulin homology

Query Match 42.4%; Score 1203; DB 2; Length 374;
Best Local Similarity 97.8%; Pred. No. 1.5e-73;
Matches 221; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 306 RPHTCPPCPAPEALGAPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVDG 365
DB 149 KTHTCPPCPAPEALGAPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVDG 208
QY 366 VEVHNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKG 425
DB 209 VEVHNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKG 268

F:137-206/Domain: immunoglobulin homology <IM2>
F:243-310/Domain: immunoglobulin homology <IM3>
F:27-83,144-204,250-308/Distulfide bonds: #status experimental
F:103/Distulfide bonds: interchain (to light chain) #status experimental
F:109,112/Distulfide bonds: interchain (to heavy chain) #status experimental
F:180/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 42.4%; Score 1205; DB 1; Length 330;
Best Local Similarity 73.7%; Pred. No. 9.5e-74;
Matches 235; Conservative 10; Mismatches 30; Indels 44; Gaps 5;

QY 253 KAVTSNVASVQCDNSDKP-----VVKPKGCGPTLGDEG----- 287
DB 16 KSTSGTALGCLVKDYFPEPTVTVSNVSGALTSQVHTFPA----VLQSGLSVLSVWTV 71
QY 288 -----DITDLY----DYYPEDTEGDKVRP-----HTCCPPCPAPEALGAPSVFLFPKPK 332
DB 72 PSSSLGTQYICNVNPKPNTKVDKVKPKCDKTHCCPPAPEALGAPSVFLFPKPK 131
QY 333 DTLMSRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTKPREQYNSTYRVVSVLTV 392
DB 132 DTLMSRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTKPREQYNSTYRVVSVLTV 191
QY 393 LHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRREMTKNQVSLTCL 452
DB 192 LHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCL 251
QY 453 VKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFPLYSKLTVDKSRWQQGNVFCVM 512
DB 252 VKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFPLYSKLTVDKSRWQQGNVFCVM 311
QY 513 HEALHNHYTQKSLSLSPGK 531
DB 312 HEALHNHYTQKSLSLSPGK 330

RESULT 3
Ig heavy chain V region precursor - human
C:Species: Homo sapiens (man)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 01-Dec-2000
C:Accession: S69339; S72664
R:Khamlichi, A.A.; Aucouturier, P.; Preud'homme, J.L.; Cogne, M.
Eur. J. Biochem. 229, 54-60, 1995
A:Title: Structure of abnormal heavy chains in human heavy-chain-deposition disease.
A:Reference number: S69339; MUID:95262687; PMID:7744049
A:Accession: S69339
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-374 <KHA>
A:Cross-references: EMBL:X81695
R:Khamlichi, A.A.
submitted to the EMBL Data Library, September 1994
A:Reference number: S72664
A:Accession: S72664
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-140, 'C', 142-374 <KH2>
A:Cross-references: EMBL:X81695
C:Superfamily: immunoglobulin C region; immunoglobulin homology

Query Match 42.4%; Score 1203; DB 2; Length 374;
Best Local Similarity 97.8%; Pred. No. 1.5e-73;
Matches 221; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 306 RPHTCPPCPAPEALGAPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVDG 365
DB 149 KTHTCPPCPAPEALGAPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVDG 208
QY 366 VEVHNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKG 425
DB 209 VEVHNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKG 268

QY 426 QPREQVYTLPPSRREMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSD 485
DB 269 QPREQVYTLPPSRREMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSD 328
QY 486 GSFFLYSKLTVDKSRWQQGNVFCVMHEALHNHYTQKSLSLSPGK 531
DB 329 GSFFLYSKLTVDKSRWQQGNVFCVMHEALHNHYTQKSLSLSPGK 374

RESULT 4
S31866
Ig gamma-1 chain C region - synthetic
C:Species: synthetic
A:Note: Homo sapiens (man) gene engineered and expressed in Escherichia coli
C:Date: 06-Jan-1995 #sequence_revision 17-Mar-1997 #text_change 19-May-2000
C:Accession: S31866
R:Pilpula, D.
submitted to the EMBL Data Library, February 1993
A:Description: Screening method for protein-protein interactions of cloned gene products.
A:Reference number: S31866
A:Accession: S31866
A:Molecule type: mRNA
A:Residues: 1-255 <PIL>
A:Cross-references: EMBL:X70421; NID:G33068; PIDN:CAA49866.1; PID:G33069
C:Keywords: immunoglobulin
F:1-22/Region: Escherichia coli outer membrane protein A precursor
F:23-255/Region: human Ig gamma-1 chain C region

Query Match 42.2%; Score 1197; DB 4; Length 255;
Best Local Similarity 96.9%; Pred. No. 2.3e-73;
Matches 219; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 306 RPHTCPPCPAPEALGAPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVDG 365
DB 30 KTHTCPPCPAPEALGAPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVDG 89
QY 366 VEVHNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKG 425
DB 90 VEVHNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKG 149
QY 426 QPREQVYTLPPSRREMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSD 485
DB 150 QPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSD 209
QY 486 GSFFLYSKLTVDKSRWQQGNVFCVMHEALHNHYTQKSLSLSPGK 531
DB 210 GSFFLYSKLTVDKSRWQQGNVFCVMHEALHNHYTQKSLSLSPGK 255

RESULT 5
PT0207
Ig gamma chain C region - chimpanzee
C:Species: Pan troglodytes (chimpanzee)
C:Date: 23-Nov-1991 #sequence_revision 23-Nov-1991 #text_change 16-Jul-1999
C:Accession: PT0207
R:Ehrlich, P.H.; Moustafa, Z.A.; Oestberg, L.
Mol. Immunol. 28, 319-322, 1991
A:Title: Nucleotide sequence of chimpanzee Fc and hinge regions.
A:Reference number: PT0207; MUID:91287716; PMID:2062315
A:Accession: PT0207
A:Molecule type: mRNA
A:Residues: 1-234 <EHR>
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: immunoglobulin
F:48-117/Domain: immunoglobulin homology <IM>

Query Match 40.6%; Score 1152; DB 2; Length 234;
Best Local Similarity 93.8%; Pred. No. 2.2e-70;
Matches 213; Conservative 3; Mismatches 5; Indels 6; Gaps 1;

QY 304 KVRP-----HTCCPPCPAPEALGAPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEV 357
DB 8 KVEPKSCDTHTCPPCAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEV 67

QY 358 KFNWYVDGVEVHNATKPREEQNSTYRVSVSLTVLHQDLNMGKEYCKVSNKALPVPIE 417
|||||
Db 68 KFNWYVDGVEVHNATKPREEQNSTYRVSVSLTVLHQDLNMGKEYCKVSNKALPAPIE 127
|||||
QY 418 KTIKAKQGPPEPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKT 477
|||||
Db 128 KTIKAKQGPPEPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKT 187
|||||
QY 478 TTPVLDSGSPFLYSLKLTVDKSRWQGNVFCSCVMHEALHNHYTOKS 524
|||||
Db 188 TTPVLDSGSPFLYSLKLTVDKSRWQGNVFCSCVMHEALHNHYTOKS 234
|||||
RESULT 6
IG gamma-2 chain C region - human
C:Species: Homo sapiens (man)
C:Date: 30-Apr-1981 #sequence_revision 13-Jun-1993 #text_change 09-Jul-2004
C:Accession: A93906; A92809; A90752; A93132; A02148
R:Ellison, J.; Hood, L.
Proc. Natl. Acad. Sci. U.S.A. 79, 1984-1988, 1982
A:Title: Linkage and sequence homology of two human immunoglobulin gamma heavy chain con
A:Reference number: A93906; MUID:82197621; PMID:6804948
A:Accession: A93906
A:Molecule type: DNA
A:Residues: 1-326 <ELL>
A:Cross-references: UNIPROT:P01859; GB:V00554; GB:J00230; NID:g32759; PIDN:CAB58438.1; F
A:Note: Lys-326 is probably removed posttranslationally
R:Wang, A.C.; Tung, E.; Fudenberg, H.H.
J. Immunol. 125, 1048-1054, 1980
A:Title: The primary structure of a human IgG2 heavy chain: genetic, evolutionary, and f
A:Reference number: A92809; MUID:81007873; PMID:6774012
A:Contents: myeloma protein Til
A:Accession: A92809
A:Molecule type: protein
A:Residues: 1-19, 'Q', 21-57, 'Z', 59, 'A', 61-193, 'D', 195-325 <WAN>
A:Note: Trp-156 is at or near the complement-binding site
R:Connell, G.E.; Parr, D.M.; Hofmann, T.
Can. J. Biochem. 57, 758-767, 1979
A:Title: The amino acid sequences of the three heavy chain constant region domains of a
A:Reference number: A90752; MUID:8001357; PMID:113060
A:Contents: myeloma protein Zie
A:Accession: A90752
A:Molecule type: protein
A:Residues: 1-24, 'E', 26-57, 'EV', 60-85, 132-171, 'ZZZ', 175, 'B', 177-193, 'D', 195-196, 'Q', 198-
A:Note: This sequence has since been revised
R:Hofmann, T.; Parr, D.M.
Mol. Immunol. 16, 923-925, 1979
A:Title: A note on the amino acid sequence of residues 381-391 of human immunoglobulin g
A:Reference number: A93132; MUID:80114419; PMID:118920
A:Contents: Zie
A:Accession: A93132
A:Molecule type: protein
A:Residues: 238-275 <HOF>
R:Hofmann, T.; Parr, D.M.
submitted to the Atlas, March 1980
A:Reference number: A94591
A:Contents: annotation: Zie, revisions to residues 25, 59, 60, and 264-268
A:Note: the revised sequence differs from that shown in having 60-Ala and in the amidati
ned
R:Milstein, C.; Frangione, B.
Biochem. J. 121, 217-225, 1971
A:Title: Disulphide bridges of the heavy chain of human immunoglobulin G2.
A:Reference number: A90253; MUID:72033500; PMID:4940472
A:Contents: annotation: myeloma protein 8a, disulfide bonds
R:Frangione, B.; Milstein, C.; Pink, J.R.L.
Nature 221, 145-148, 1969
A:Title: Structural studies of immunoglobulin G.
A:Reference number: A93157; MUID:69064124; PMID:5782707
A:Contents: annotation: 8a, disulfide bonds
C:Genetics:
A:Gene: GDB:IGHG2

A:Cross-references: GDB:119338; OMIM:147110
A:Map position: 14q32.33-14q32.33
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap
bain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F:20-85/Domain: immunoglobulin homology <IM1>
F:133-202/Domain: immunoglobulin homology <IM2>
F:239-306/Domain: immunoglobulin homology <IM3>
F:14/Disulfide bonds: interchain (to light chain) #status experimental
F:27-83,140-200,246-304/Disulfide bonds: #status experimental
F:102,103,106,109/Disulfide bonds: interchain (to heavy chain) #status experimental
F:176/Binding site: carbohydrate (Asn) #status predicted
Query Match 40.4%; Score 1148; DB 1; Length 326;
Best Local Similarity 72.2%; Pred. No. 6.4e-70;
Matches 226; Conservative 18; Mismatches 33; Indels 36; Gaps 6;
QY 253 KAVTSNVASVQCDNSDKFP-----VYKYP-----GKGCPYL----- 283
Db 16 RSTSESTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVTVVPSN 75
QY 284 -GDEGDTLDYDYEEDEGDKVRPH-----TCPPCPAPEALGAPSVLFPKPKDTLMIS 338
Db 76 FGTQTYTCNVDHKP-SNTKVDKTVKRCCKVECCPCPPVAG-PSVFLFPPKPKDTLMIS 133
QY 339 RTPVETCVVDVSHEDPEVFNWVDCGVVEVHNATKPREQNSTYRVSVSLTVLHQDLN 398
Db 134 RTPVETCVVDVSHEDPEVFNWVDCGVVEVHNATKPREQNSTYRVSVSLTVLHQDLN 193
QY 399 NGKEYCKVSNKALPVPPIETKISKAKQGPPEPQVYTLPPSREEMTKNQVSLTCLVKGFYP 458
Db 194 NGKEYCKVSNKGLPAPIETKISKAKQGPPEPQVYTLPPSREEMTKNQVSLTCLVKGFYP 253
QY 459 SDIAVEWESNGQPENNYKTTPPVLDSGSPFLYSLKLTVDKSRWQGNVFCSCVMHEALHN 518
Db 254 SDIAVEWESNGQPENNYKTTPPVLDSGSPFLYSLKLTVDKSRWQGNVFCSCVMHEALHN 313
QY 519 HYTKSLSLSPGK 531
Db 314 HYTKSLSLSPGK 326
RESULT 7
Ig gamma-3 chain C region (allotype G3m(b)) - human
C:Species: Homo sapiens (man)
C:Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 23-Jul-1999
C:Accession: A23511
R:Huck, S.; Fort, P.; Crawford, D.H.; Lefranc, M.P.; Lefranc, G.
Nucleic Acids Res. 14, 1779-1789, 1986
A:Title: Sequence of a human immunoglobulin gamma 3 heavy chain constant region gene: co
A:Reference number: A23511; MUID:86148507; PMID:3081877
A:Accession: A23511
A:Molecule type: DNA
A:Residues: 1-377 <HUC>
A:Cross-references: GB:X03604; GB:M12958; NID:g33070; PIDN:CAA27268.1; PID:g577056
C:Genetics:
A:Gene: GDB:IGHG3
A:Cross-references: GDB:119339; OMIM:147120
A:Map position: 14q32.33-14q32.33
A:Introns: 98/3; 115/3; 130/3; 145/3; 160/3; 270/3
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: immunoglobulin
F:20-85/Domain: immunoglobulin homology <IMM>
Query Match 40.3%; Score 1145; DB 2; Length 377;
Best Local Similarity 64.4%; Pred. No. 1.2e-69;
Matches 230; Conservative 17; Mismatches 56; Indels 54; Gaps 7;
QY 194 LQGENSLY-----TTPKGFEGSHLLPFAFLHGNPWLNCCEILYFRWLQDAENVVVK 247
Db 56 VLOSSGLYSLSSVTVVPSLSLGT-----QTYTCN-----VNHKPSNTKVDK 96

Biochemistry 19, 4304-4308, 1980
 A>Title: Primary structure of human gamma3 immunoglobulin deletion mutant: gamma3 heavy-
 A:Reference number: A90442; MUID:81021548; PMID:6774747
 A:Contents: heavy chain disease protein Wis
 A:Accession: A90442
 A:Molecule type: protein
 A:Residues: 1-289 <FRA>
 A>Note: the molecule is a dimer linked by 12 disulfide bonds; it has an extra interchain
 A>Note: this protein lacks most of the V region and all of the CH1 region. Residue 12 co
 A>Note: the sequence of residues 42-76 was taken from the reference that follows
 R:Michaelson, T.E.; Frangione, B.; Franklin, E.C.
 J. Biol. Chem. 252, 883-889, 1977
 A>Title: Primary structure of the 'hinge' region of human IgG3. Probable quadruplication
 A:Reference number: A92219; MUID:71118561; PMID:402363
 A:Contents: normal gamma-3 chains, sequence corresponding to residues 12-97 of protein W
 A:Accession: A92219
 A:Molecule type: protein
 A:Residues: 12-97 <MIC>
 A>Note: the hinge region in gamma-3 chains is about four times as long as in other gamma
 idue segment (12-28)
 A>Note: cysteines at positions 24, 27, 33, 39, 42, 48, 54, 57, 63, 69, and 72 form inter
 R:Wolfenstein-Todel, C.; Frangione, B.; Prelli, F.; Franklin, E.C.
 Biochem. Biophys. Res. Commun. 71, 907-914, 1976
 A>Title: The amino acid sequence of "heavy chain disease" protein ZUC. Structure of the
 A:Reference number: A90198; MUID:77021516; PMID:823945
 A:Contents: heavy chain disease protein Zuc, partial sequence corresponding to residues
 A:Accession: A90198
 A:Molecule type: protein
 A:Residues: 59-125, 'EB', 128-226, 228-289 <WOL>
 A>Note: this protein lacks most of the V region, all of the CH1 region, and part of the
 R:Alexander, A.; Steimetz, M.; Barritault, D.; Frangione, B.; Franklin, E.C.; Hood, L.;
 Proc. Natl. Acad. Sci. U.S.A. 79, 3260-3264, 1982
 A>Title: gamma heavy chain disease in man: cDNA sequence supports partial gene deletion
 A:Reference number: A93915; MUID:82247835; PMID:6808505
 A:Contents: heavy chain disease protein Omm
 A:Accession: A93915
 A:Molecule type: mRNA
 A:Residues: 12-70; 72-114; 116-125, 'E', 127-133, 'L', 135-136, 'E', 138, 'Y', 140-154, 'D', 156-157
 A>Note: a carboxyl-terminal Lys is removed posttranslationally
 A>Note: this sequence may represent an allelic form or another gamma chain subclass
 C:Comment: The heavy chain disease protein Wis is shown.
 C:Genetics:
 A:Gene: GDB:IGHG3
 A:Cross-references: GDB:119339; OMIM:147120
 A:Map position: 14q32.33-14q32.33
 C:Superfamily: immunoglobulin C region; immunoglobulin homology
 C:Keywords: duplication; glycoprotein; immunoglobulin; pyroglutamic acid
 F:203-270/Domain: immunoglobulin homology <IMW>
 F:1/Modified site: pyroglutamic acid (Gln) #status experimental
 F:6,140/Binding site: carboxylate (Asn) (covalent) #status experimental

Query Match 39.3%; Score 1116.5; DB 1; Length 289;
 Best Local Similarity 78.7%; Pred. No. 7.2e-68;
 Matches 210; Conservative 11; Mismatches 21; Indels 25; Gaps 2;

Qy 264 CDNSDKFVYKPGKCTPLGDEGDTLDYVYPEDETEGDKVRPHTCPCPAPALGAPS 323
 |||
 Db 48 CDTPPPCRCPEP-KSCDT-----PPPCRCPAPALGAPS 82
 Qy 324 VFLEPPKPKOTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNST 383
 |||
 Db 83 VFLEPPKPKOTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNST 142
 Qy 384 YRVSVLTVLHODWLNKGEYKCKVSNKALPVPKEITISKAKGPQEPQVYTLPSREMT 443
 :|||
 Db 143 FRVSVLTVLHQNLDGKEYKCKVSNKALPAPIEKTISKTKGQPREPQVYTLPSREMT 202
 Qy 444 KNQVSLTCLVKGFPSPDIAVENSNGQPENNYKTTPVLDSDGSFFLYSKLTVDKSRWQQ 503
 |||
 Db 203 KNQVSLTCLVKGFPSPDIAVENSNGQPENNYNTTPMLDSDGSFFLYSKLTVDKSRWQQ 262
 Qy 504 GNVFSCSVHHEALHNHYTQKSLSLSPG 530
 |||

Db 263 GNVFSCSVHHEALHNHYTQKSLSLSPG 289

RESULT 11

GHRB

Ig gamma chain C region - rabbit
 C:Species: Oryctolagus cuniculus (domestic rabbit)
 C>Date: 24-Apr-1984 #sequence revision 15-Nov-1984 #text_change 09-Jul-2004
 C:Accession: A91749; A90290; A93928; A90445; A94416; A02161
 R:Bernstein, K.E.; Alexander, C.B.; Mage, R.G.
 Immunogenetics 18, 387-397, 1983
 A>Title: Nucleotide sequence of a rabbit IgG heavy chain from the recombinant F-I haplot
 A:Reference number: A91749; MUID:84030930; PMID:6313520
 A:Accession: A91749
 A:Molecule type: mRNA
 A:Residues: 1-323 <BER>
 A:Cross-references: UNIPROT:P01870
 A>Note: this sequence has the d12 allotypic marker, 104-Thr, and the e14 marker, 185-Thr
 R:Pratt, D.M.; Mole, L.E.
 Biochem. J. 151, 337-349, 1975
 A>Title: Sequence studies on the constant region of the Fd sections of rabbit immunoglob
 A:Reference number: A90290; MUID:76135469; PMID:1243651
 A:Accession: A90290
 A:Molecule type: protein
 A:Residues: 1-47, 'E', 49-71, 'PV', 72-128 <PRA>
 R:Martens, C.L.; Moore, K.W.; Steinmetz, M.; Hood, L.; Knight, K.L.
 Proc. Natl. Acad. Sci. U.S.A. 79, 6018-6022, 1982
 A>Title: Heavy chain genes of rabbit IgG; isolation of a cDNA encoding gamma heavy chain
 A:Reference number: A93928; MUID:83299917; PMID:6193512
 A:Accession: A93928
 A:Molecule type: mRNA
 A:Residues: 88-103, 'M', 105-143, 'E', 145-184, 'A', 186, 'E', 188-266 <MAR>
 A:Cross-references: GB:M16426; NID:9165111; PID:AA31289.1; PID:9165112
 A>Note: this sequence has the d11 allotypic marker, 104-Met, and the e15 allotypic mark
 R:Frutcher, R.G.; Jackson, S.A.; Mole, L.E.; Porter, R.R.
 Biochem. J. 116, 249-259, 1970
 A>Title: Sequence studies of the Fd section of the heavy chain of rabbit immunoglobulin
 A:Reference number: A90245; MUID:70110015; PMID:5461106
 A:Accession: A90245
 A:Molecule type: protein
 A:Residues: 132-143, 'E', 145-161 <FRU>
 R:Hill, R.L.; Lebovitz, H.E.; Fellows Jr., R.E.; Delaney, R.
 in Gamma Globulins, Nobel Symp. 3, Killander, J., ed., pp.109-127, Almqvist and Wiksell,
 A:Reference number: A94416
 A:Accession: A94416
 A:Molecule type: protein
 A:Residues: 129-131; 155-172, 'D', 174-184, 'A', 186, 'E', 188-200, 'D', 202-217, 'E', 219-232, 'Q',
 A>Note: this has the e15 allotypic marker, 185-Ala
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap
 hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la
 C:Superfamily: immunoglobulin C region; immunoglobulin homology
 C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
 F:20-82/Domain: immunoglobulin homology <IMI>
 F:130-199/Domain: immunoglobulin homology <IM2>
 F:236-303/Domain: immunoglobulin homology <IM3>
 F:173/Binding site: carboxylate (Asn) (covalent) #status predicted

Query Match 32.4%; Score 921; DB 1; Length 323;
 Best Local Similarity 61.1%; Pred. No. 1.2e-54;
 Matches 173; Conservative 35; Mismatches 45; Indels 30; Gaps 3;

Qy 252 VKAVTSNVASVQCDNSDKFPVYKPGKCTPLGDEGDTLDYVYPEDETEGDK-VRPHTC 310
 :|||
 Db 68 VVSVTSNQPVTCNVA-----HPATNTKVDKTVAPSTC 100
 Qy 311 --PPCPAPEALGAPSVLEFPFKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEV 368
 |||
 Db 101 SKPTCPPELLGGPSVIFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFTWYINNEQV 160
 Qy 369 HNAKTKPREQYNSTIRRVSVLTVLHODWLNKGEYKCKVSNKALPVPKEITISKAKQPR 428
 :|||
 Db 161 RTARPPPLREQQFNSTIRRVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTISKAKQPL 220

```

QY 429 EPOVYTLPPSREEMTKNOVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLDSDGSF 488
Db 221 EPKYITMGPPRELSRSVSLTCLMNGFYPSDIAVWESNGKAEDNYKTTPPAVLDSGSY 280

QY 489 FLYSKLTVDKSRWQGVNFVSCSVNHEALHNYTKSLSPGK 531
Db 281 FLYNKLSTVPTSEWQGDVFTCSVNHEALHNYTKSLSPGK 323

RESULT 12
I47160
Ig gamma 2b chain constant region - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
C:Accession: I47160
R:Kacskovics, I.; Sun, J.; Butler, J.E.
A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of a
A:Reference number: I47158; MUID:95015845; PMID:7930579
A:Accession: I47160
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-328 <KAC>
A:Cross-references: EMBL:U03780; NID:G4333125; PIDN:AA52218.1; PID:G4333126
C:Genetics:
A:Gene: IgG2b
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:133-202/Domain: immunoglobulin homology <IMM>

Query Match 32.2%; Score 913; DB 2; Length 328;
Best Local Similarity 71.6%; Pred. No. 4.2e-54;
Matches 166; Conservative 31; Mismatches 31; Indels 4; Gaps 3;

QY 302 GDKVRPHHTPCPCAPALGAPSVFLPPPKKDTLMISRTPEVTCVVVDVSHEDPEVKFNW 361
Db 99 GTKTKP-PCPICPACESPG-PSVFIPLPPPKKDTLMISRTPEVTCVVVDVSHEDPEVKFNW 156

QY 362 YVDGVEVHNKTPREEQVNSTYRVVSVLTVLHODWLNKGEYKCKYKSNKALPVPPIETIS 421
Db 157 YVDGVEVHTAQRPEEQNFNFTYRVVSVLPIQHODWLNKGEYKCKYKSNKALPVPPIETIS 216

QY 422 KAGQPREPQVYTLPPSREEMTKNOVSLTCLVKGFYPSDIAVWESNGQ--PENNYKTTP 479
Db 217 KAGQPREPQVYTLPPHAEELSRKSVITCLVIGFYPPDIDVWQNGQPEPEGNVYRTTP 276

QY 480 PVLDSGSPFLYSKLTVDKSRWQGVNFVSCSVNHEALHNYTKSLSPGK 531
Db 277 PQQVDVGTLYSKFSVDKASWQGGIFQCAVMHEALHNYTKSLSPGK 328

RESULT 13
I47159
Ig gamma 2a chain constant region - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
C:Accession: I47159
R:Kacskovics, I.; Sun, J.; Butler, J.E.
A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of a
A:Reference number: I47158; MUID:95015845; PMID:7930579
A:Accession: I47159
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-328 <KAC>
A:Cross-references: EMBL:U03779; NID:G4333123; PIDN:AA52217.1; PID:G4333124
C:Genetics:
A:Gene: IgG2a
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:133-202/Domain: immunoglobulin homology <IMM>

Query Match 32.2%; Score 913; DB 2; Length 328;
Best Local Similarity 71.6%; Pred. No. 4.2e-54;
Matches 166; Conservative 31; Mismatches 31; Indels 4; Gaps 3;

```

```

QY 302 GDKVRPHHTPCPCAPALGAPSVFLPPPKKDTLMISRTPEVTCVVVDVSHEDPEVKFNW 361
Db 99 GTKTKP-PCPICPACESPG-PSVFIPLPPPKKDTLMISRTPEVTCVVVDVSHEDPEVKFNW 156

QY 362 YVDGVEVHNKTPREEQVNSTYRVVSVLTVLHODWLNKGEYKCKYKSNKALPVPPIETIS 421
Db 157 YVDGVEVHTAQRPEEQNFNFTYRVVSVLPIQHODWLNKGEYKCKYKSNKALPVPPIETIS 216

QY 422 KAGQPREPQVYTLPPSREEMTKNOVSLTCLVKGFYPSDIAVWESNGQ--PENNYKTTP 479
Db 217 KAGQPREPQVYTLPPHAEELSRKSVITCLVIGFYPPDIDVWQNGQPEPEGNVYRTTP 276

QY 480 PVLDSGSPFLYSKLTVDKSRWQGVNFVSCSVNHEALHNYTKSLSPGK 531
Db 277 PQQVDVGTLYSKFSVDKASWQGGIFQCAVMHEALHNYTKSLSPGK 328

RESULT 14
I47162
Ig gamma 4 chain constant region - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
C:Accession: I47162
R:Kacskovics, I.; Sun, J.; Butler, J.E.
A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of a
A:Reference number: I47158; MUID:95015845; PMID:7930579
A:Accession: I47162
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-277 <KAC>
A:Cross-references: EMBL:U03782; NID:G4333129; PIDN:AA52220.1; PID:G4333130
C:Genetics:
A:Gene: IgG4
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:82-151/Domain: immunoglobulin homology <IMM>

Query Match 31.9%; Score 906; DB 2; Length 277;
Best Local Similarity 71.1%; Pred. No. 9.9e-54;
Matches 165; Conservative 30; Mismatches 33; Indels 4; Gaps 3;

QY 302 GDKVRPHHTPCPCAPALGAPSVFLPPPKKDTLMISRTPEVTCVVVDVSHEDPEVKFNW 361
Db 48 GTKTKP-PCPICPACESPG-PSAFIPLPPPKKDTLMISRTPEVTCVVVDVSHEDPEVKFNW 105

QY 362 YVDGVEVHNKTPREEQVNSTYRVVSVLTVLHODWLNKGEYKCKYKSNKALPVPPIETIS 421
Db 106 YVDGVEVHTAQRPEEQNFNFTYRVVSVLPIQHODWLNKGEYKCKYKSNKALPVPPIETIS 165

QY 422 KAGQPREPQVYTLPPSREEMTKNOVSLTCLVKGFYPSDIAVWESNGQ--PENNYKTTP 479
Db 166 KAGQPREPQVYTLPPHAEELSRKSVITCLVIGFYPPDIDVWQNGQPEPEGNVYRTTP 225

QY 480 PVLDSGSPFLYSKLTVDKSRWQGVNFVSCSVNHEALHNYTKSLSPGK 531
Db 226 PQQVDVGTLYSKLAVDKASWQGGIFQCAVMHEALHNYTKSLSPGK 277

RESULT 15
I47158
Ig gamma 1 chain constant region - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
C:Accession: I47158
R:Kacskovics, I.; Sun, J.; Butler, J.E.
A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of a
A:Reference number: I47158; MUID:95015845; PMID:7930579
A:Accession: I47158
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-328 <KAC>

```

Search completed: May 24, 2005, 06:09:00
Job time : 44 secs